DOCUMENT RESUME

ED 092 169 IR 000 733

TITLE	Facts and Futures; What's Happening Now in Computing for Higher Education. Annual Proceedings of the EDUCOM Fall Conference (9th, Princeton, New Jersey, October 9-11, 1973).
INSTITUTION	Interuniversity Communications Council (EDUCOM), Princeton, N. J.
PUB DATE	74
NOTE	361p.
AVAILABLE FROM	EDUCOM, P. O. Box 364, Princeton, New Jersey 08540 ($9.00)
EDRS PRICE	MF-$0.75 HC-$17.40 PLUS POSTAGE
DESCRIPTORS	Computer Assisted Instruction; Computer Oriented Programs; *Computers; Computer Science; *Conference Reports; Cost Effectiveness; Data Bases; *Higher Education; Information Centers; *Information Science; *Information Systems; Management Information Systems; Professional Personnel; Research; Statewide Planning
IDENTIFIERS	Compatibility; EDUCOM; Transportability

ABSTRACT

The ninth annual EDUCOM conference developed its theme along four major lines: computing for research; computing for instruction; management information systems; and the allocation of computing resources. Papers in this volume address primarily the organizational and political considerations (Parts II, III, IV, and V), and technological/economic issues (Part VI) relevant to the utilization of computing and networking in higher education. (WCM)

ED 092169

FACTS and FUTURES

WHAT'S HAPPENING NOW IN COMPUTING FOR HIGHER EDUCATION

PROCEEDINGS

of the

EDUCOM

FALL CONFERENCE

October 9, 10, 11, 1973
Princeton, New Jersey

EDUCOM, The Interuniversity Communications Council, Inc.

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

EDUCOM

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER"

Library of Congress Catalog Card Number: 74-79222

Further information and additional copies of this book may be obtained from EDUCOM, P.O. Box 364, Princeton, New Jersey 08540. The price per copy is $5.00 for members and $9.00 to nonmembers of EDUCOM.

©1974 EDUCOM, The Interuniversity Communications Council, Inc.
Printed in the United States of America

Contents

Preface 1

Introduction 3

PART I. KEYNOTE

1. Administration of Computer Resources
GENE F. FRANKLIN 7

2. Management Information Systems
DONALD C. CARROLL 13

3. Computers in Instruction
CHARLES H. WARLICK 18

4. Computing in Research
ROBERT L. ASHENHURST 28

PART II. ADMINISTRATION OF COMPUTER RESOURCES

5. Administration of Computer Resources: An Overview
JAMES POAGE 37

6. Alternatives In Computer Administration

The Games Universities Play
JOHN E. SKELTON 40

The Information Utility in a University Environment
PAUL STEPHAN 46

The Future of University Computing Centers
WILLIAM KEHL 52

7. Considerations of Networking

Network Management for Expanded Resource Sharing
DENNIS W. FIFE 55

An Information Dissemination Network Model
GLENN O. WARE and
JOHN H. SCHUENEMEYER 62

The Economics of University Computer Networking
D. A. DUNN 68

8. Statewide Plans for Computing

Introduction
E. REX KRUEGER 73

Educational Information Services, Inc.
GILBERT G. MOSER 75

Critique of the New Jersey Plan
FREEMAN HOLMER 78

Critique of the New Jersey Plan
DAVID McINTYRE 81

Computer Services in the Oregon
Department of Higher Education
MICHAEL A. JENNINGS 83

Critique of Oregon Network Plan
EUGENE KEENOY 100

The Florida Regional Computing
Center Plan
JAMES MORGAN 104

Critique of Florida Computing Plan
GEORGE STRUBLE 109

Computer Activities in Illinois
TIMOTHY GRIFFIN 111

What is IECCS?
RONALD BRADY 112

Statewide Networks for Computing
CHARLES J. MOSMANN 115

PART III. MANAGEMENT INFORMATION SYSTEMS

9. An Overview
HAROLD WAKEFIELD 121

10. MIS Case Studies

The Effects of Management Information
Systems in Colleges & Universities: A
Perspective from the Statewide Level
W. K. BOUTWELL 124

MIS: The Ugly Duckling
JOHN W. GWYNN 129

Project FIND
JOHN S. McGEACHIE 137

Evolutionary Tendencies of MIS at
the University of Houston
RALPH C. McKAY 146

11. **Exporting and Importing MIS**

Exportability and the In-house MIS
CHERYL M. TRAVER 153

On-line Administrative Information
Systems: A Case Study
PAUL W. SIRE 156

Importing MIS Components:
A Practitioner-Manager Perspective
HENRY G. VAUGHAN 164

One University's View of
Data Management Systems
KENNETH E. SHOSTACK 174

PART IV. COMPUTING IN INSTRUCTION

12. **Computing in Instruction:**
An Overview
WILLIAM ATCHISON 187

13. **Instructional Computing Case Studies**

Computers in Instruction
KENNETH L. BOWLES 190

Free-Access Computing at Dartmouth
ARTHUR W. LUEHRMANN 193

Computing Services at Chicago
FRED H. HARRIS 198

Trends in Instructional Use of Computers
GERARD P. WEEG 206

14. Transportability of Instructional Systems

Networking Challenges: The User's Viewpoint
THOMAS N. PYKE, JR. and
ROBERT P. BLANC 211

CAI Programs for Multi-University Use
KEITH A. HALL 218

Scientific and Technical Information
Transfer for Education (STITE)
PRANAS ZUNDE 229

Chemical Engineering in Structural Packages
WARREN D. SEIDER 231

15. Software Compatibility in the Social Sciences

Performance Differences in Social
Science Statistical Software
WILLIAM D. SLYSZ 235

Software Compatibility in
Social Science Computing
JUDITH ROWE 244

Reducing Incompatibilities in Social
Science Software and Data: A Report
on the SIGSOC Workshop
RONALD E. ANDERSON 246

PART V. COMPUTERS IN RESEARCH

16. **Computers in Research: An Overview**
DOUWE B. YNTEMA 257

17. **Information Services for Research**

New Channels of Distribution in the Information Industry
WILLIAM R. NUGENT 260

NASIC: A Regional Experiment in the Brokerage of Information Services
DAVID M. WAX and
R. D. MORRISON, JR. 268

The Ohio State University Mechanized Information Center
HUGH C. ATKINSON 274

Data Bases and Libraries
G. A. HARRER

18. **Resource Sharing for Libraries**

Major Trends in Library Computerization
RICHARD DE GENNARO 282

Cooperative Automation for Libraries
RUSSELL SHANK 286

19. **Discipline Oriented Applications**

Hierarchical Computing for Chemistry
JOHN CORNELIUS 289

Computer-Based Medical Records
ELEMER GABRIELI 293

User Issues in Medical Record Automation
ROBERT M. STRONG 299

Facilities for Data Manipulation in the Humanities
MICHAEL E. LESK 305

Steps Toward Making Literary Texts Available
STEPHEN V. F. WAITE 309

PART VI. A BROADER VIEW

20. Data Transmission Services

Regulation of Computer Communications
ROBERT P. BIGELOW, ESQUIRE 317

Bell System Services for Digital Data Transmission
CARL F. STUEHRK 322

PCI's VANLINE Service
LEE B. TALBERT 327

Appendix A – FIND System Commands 333

Appendix B – FIND Available Data Bases (December 1973) 335

Appendix C – Resident File OMIC, 1973 337

Appendix D – Sample Queries of University of Vermont Data Bases 341

Appendix E – Agenda SIGSOC Working Session on COMPATIBILITY 345

Appendix F – Conference Participants 347

Preface

A sense of purpose and an underlying confidence that things long talked of were beginning to happen characterized the Ninth Annual EDUCOM Fall Conference, *Facts and Futures.* After several decades of experimentation with modes of delivering computing services to higher education, colleges and universities now view computing resources in quite a different light than when they were first introduced. The use of computers for research is taken for granted and one now talks seriously about resource sharing, management and organization rather than solely about the technical aspects of using the computer. Resource sharing can be discussed only because resources which are worth sharing have been created. The related questions of management and organization must be considered because computing is now used in all functions of the university from instruction and research to administration. Although computing in higher education is still in a state of transition, the use of computers for educational purposes has in many ways come of age.

Extending the theme of resource sharing for computing in higher education which has run through previous EDUCOM meetings, the Fall 1973 Conference, *Facts and Futures: What's Happening Now in Computing for Higher Education*, developed this theme along four major lines: Computing for Research; Computing for Instruction; Management Information Systems; and the Allocation of Computing Resources. Spanning the wide range of EDUCOM members, this conference brought together representatives of many disciplines from universities, colleges, junior colleges, and educational service organizations to share their experiences and their views of what is happening now in computing for higher education. The attendance of 340 surpassed that of all previous EDUCOM meetings and included participants from Italy, Africa, Canada and Puerto Rico as well as both coasts of the continental USA.

Keynote addresses and papers presented in panels and workshops have been edited and collected in the following pages. Further information concerning any of the systems or applications described in these chapters can best be obtained by writing directly to the author of the presentation. The names and address of all conference participants are listed in the back of this volume.

On behalf of all the conferees, I want to extend sincere thanks to the Conference Chairman, Joe B. Wyatt, for the stimulating and comprehensive program which he developed for the conference. Special thanks are also due to Charles Mosmann and James Poage who assisted in arranging the panels on statewide computing planning, and to Julian Bigelow for providing a most pleasant evening of reminiscence on the early days of computing.

Henry Chauncey

Introduction

While computing is in a state of transition for higher education, some portents for the future are obvious. Networking as a technology and a delivery mechanism for computing services is here to stay. The real question is what role will networking play relative to colleges and universities in this country? In 1973, point-to-point digital communications facilities, star networks, and the ARPA prototype of packet-switched networks, have become established forms of computer communication. Institutions must now consider the appropriate balance between information and computing services provided on campus and those obtained off campus. Four sets of issues are prominent:

- Organizational considerations
- Political considerations
- Intellectual Issues
- Technological/Economic Issues

Papers in this volume address primarily the organizational considerations, political considerations, and technological/economic issues relevant to the utilization of computing and networking in higher education. Following a general review of these issues in relation to each of the main topic areas of the conference, presentations in Parts II, III, IV and V address the issues of organizational considerations and political considerations for computing and networking in higher education.

In Chapter 8, Statewide Plans for Computing in three states are presented and criticized. In order to elicit pointed and relevant critiques, computer center directors and college and university administrators from one state were invited to analyze the statewide computing plans of another state. Statewide plans are presented for New Jersey, Oregon, Illinois and Florida by representatives of the central agency for higher education in each state and critiques are given by computer center directors and administrators from other states. An overview of statewide activities for academic computer planning in the American States and Canadian Provinces is given in the closing paper for this chapter by Charles J. Mosmann.

Papers in Part III, Chapters 9 through 11, focus on what appears to be a new trend in management information systems for colleges and universities: The cost and effort required to develop custom management information systems has become so great that many colleges and universities are now considering importing MIS from other institutions or from information system companies. Examples of activities at both exporting and importing institutions outline for the reader the difficulties and some successful experiences.

The technological/economic issues are addressed in part VI. Each author contributing to this section has been asked to discuss: the transmission of information; how one interfaces with transmissions; and how interfacing can be facilitated. In addition, Carl Stuerhk and Lee Talbert outline plans for data transmission services of ATT and Packet Communications, Inc. which have been developed in response to the increasing availability of data transmission services through microwave and packet switching techniques.

Colleges and universities are considering the proper balance between in-house development and outside purchase of software packages for administration, instruction and research. The cost effectiveness of obtaining specific computing services via networks or importation should become more apparent as more institutions follow this trend.

Joe B. Wyatt
Conference Chairman

PART I

KEYNOTE

5/6

Chapter 1

Administration of Computer Resources

by Gene F. Franklin
Stanford University

I've been inspired to choose my text as biblical from the book of Matthew, Fifth Chapter, Verses 11 and 11a. The first verse is for directors of computer centers: "Blessed are ye when men shall revile you falsely and shall say all manner of evil against you falsely." Verse 11a, for those expected to be relieved of all their computer difficulties reads: "Blessed is he who expects nothing for he shall not be disappointed." This paper will describe briefly some of the background and recent decisions affecting computing at Stanford and to note a few lessons that might be learned from the Stanford experience.

COMPUTING SERVICES BEFORE 1973

A few years ago Stanford had three computer organizations and five general computer facilities, in addition to the special facilities like the Artificial Intelligence Laboratory that John McCarthy runs in the Department of Computer Science. Stanford owns and operates a substantial hospital in conjunction with the medical school. For patient billing, general accounts and other administrative functions, the hospital had a data processing department which has reported to the hospital financial officer. In the hospital data processing department an IBM 40 was replaced by an IBM 370/135 operating under the disk operating system DAS with a special monitor called the Shared Hospital Accounting System, or SHAS. In this facility a negligible systems programming staff, a sizeable group of applications programmers and a data control section ran census updates and posted chargeable items daily on a machine generally tuned for the particular financial support of a university hospital.

The university office of business and finance uses an IBM 370/145 (also a disk operating system) running COBOL programs in support of finance and student services such as registration and student aid. In addition, this facility supports the development office in a major campaign seeking $300 million for Stanford University, provides accounting support for faculty engaged in research under federal government contracts and privately sponsored contracts, and supports academic planning. Finally, this facility ran a sizeable billing service for the physicians on the faculty of the medical school who operate the Stanford University medical clinic. This facility, like the hospital computer facility, was tuned to the requirements of a particular group of users and rendered a particular style of service. Somewhat unusual for such a center was the fact that a significant advanced development program in computer support for University administration (INFO) was being conducted in the Management Systems Office under Michael Roberts, Director, and John Gwynn.

In the area of academic computing in support of research and instruction, the Stanford Computation Center – which reported to the provost who is the chief academic officer at Stanford – operated three computer facilities. The most specialized of these, ACME, used a 360/50 with substantial support from the NIH to provide computer support for medical research. Implemented by Gio Wiederhold and his staff, the single language, terminal-oriented facility aimed to be so easy to use that the physician-researcher would do most of his own programming and maintain control over and have access to his own data and data files. The facility had extensive provision for data collection and plotting including a centralized analog to digital conversion facility, and a graphic support software for a variety of devices including plotters, CRT displays, and the like. The specialized machine configuration for ACME included a minimum of high speed memory, but two million bytes of low speed core which was rather unusual five years ago in a model 50. The language processor, PL/ACME, was a re-entrant, core resonant, variant of PL-1 running under IBM's operating system OS MVT. A small systems staff which also ran a consultant service for physicians and users, maintained the processor and the locally designed software.

Also managed by the Stanford Computation Center was the Stanford Linear Accelerator Center (SLAC) computer facility operated by Stanford for the Atomic Energy Commission as a national research facility in physics. The SLAC IBM 360/91 operated mainly FORTRAN as a number cruncher for a relatively small number of highly sophisticated users. A substantial systems programming staff, who were charged to be at the forefront of scientific computing, also maintained an extensive tape library of user data and wrote a text editor and remote job entry system, WYLBUR, which is now available at a number of computing centers around the country.

The last of Stanford's five computer facilities was the Campus facility which provided computing for anyone who did not qualify for service from the other four. The charter for the campus facility was to provide the greatest variety of services with the maximum flexibility, fast turnaround while being reliable and cost effective. Serving about 5,000 users and running between 10,000 and 12,500 jobs a week, the campus computer center 360/67, operating under OS MFT, was dynamically switched between real and virtual and back to real storage again, under systems modifications designed at Stanford about six years ago when IBM's announced timesharing system failed to be efficient enough for us to be able to afford it.

The campus computer center staff also developed a "high speed" or "quick partition" service primarily for short jobs like student projects and homework. Although the timesharing system was available, the pricing policy and machine effectiveness of timesharing was such that the majority of jobs were, and still are, submitted through the card readers. This facility has a large staff of systems programmers dedicated to system maintenance and enhancement, and a large user services staff to provide a wide range of courses and consulting to help an extremely diverse user community make the best use of the system.

In summary, computing service at Stanford two years ago included five facilities: the hospital, university administration, medical school, SLAC and general academic facilities. Four appealed to particular user communities, and only one served a truly diverse group of users.

ORGANIZATIONAL CHANGE

The direction of change in organization of these facilities should be noted. Forces influencing this change in organization of the computing facilities were technological, programmatic, financial and managerial. While hospital data processing was doing patient billing, the university administrative facility was giving computing support to the medical clinics. There was a substantial amount of overlap in these functions. Many hospital patients were also clinic patients and many physicians who ran the clinics also had admitted patients. Consequently, the hospital systems staff embarked on a program to design a unified patient accounting system to handle all the bills and records, indicating that they needed to increase the size and, to some extent, the character of their computing.

On the other hand, many of the physicians in the medical school who had developed systems on the ACME research computer, now sought computing services in a production rather than research and development mode. The hospital data processing department felt that many of these production services should be part of the centralized computing services, and viewed with some concern the initiation of independent systems in pathology where an XDS Sigma 3 was installed to run the pathology laboratory and maintain laboratory records. It was apparent that a large amount of data communication between the laboratory automation system and the financial billing system would be necessary. Similarly, a research program in the department of medicine on drug interaction initiated by Dr. Stanley Cohen needed to build a data base dependent in part upon drug administration data from all the patients at the Stanford Medical Center. He, too, had instituted an independent computer support based upon a DEC PDP 11/45, and had begun to build a data base for his drug interaction studies. Underlying all of this was the concern at Stanford medical center, as in many other centers, about a growing trend toward on-line hospital information systems. And although it was far from clear what kind of system Stanford would introduce, it was evident that a 360/40 or a 370/135 operating under a financial officer was an unlikely place for development of a university medical center and hospital system. In any event, the data processing department in the hopsital which had begun to design an on-line admission system using an IBM product, CICS (Customer Information Control System), to give on-line access to some of their data bases, foresaw the need to upgrade their computer facility from a 370/135 to a 370/145 and to move from a completely batch oriented environment to an on-line data base management style of computing.

Many of the same pressures, of course, were in operation at the ACME medical research facility. Both because of its design and some inherent unreliabilities and also because of its charge as a research project on advanced computing in medicine, the ACME system was not a suitable facility for production. The final irresistible force was the fact that the National Institute of Health informed the principal investigator that continuing support of the ACME facility as a service facility would be impossible.

Thus, the medical community, university hospital and medical school began an intensive study of computing needs and opportunities in the medical center. After studying the problem for some time, they proposed a central facility to serve all the medical center so that patient care data would be made available to physicians interested in primary care, to those interested in research questions and to those interested in the administration of the hospital and school. Administratively, these proposals were accorded a receptive hearing by Dr. Clayton Rich, the new dean of the medical school and vice president of the university for medical affairs. However, there were some difficulties with the funding of the center and some concern as to whether a facility in the medical center alone was the right step to take at that time. However, the need was acknowledged for a facility that would give top priority to reliability, provide on-line interactive computing including the language PL/ACME, provide extensive file management and data base support services and include provision for entering and interacting with laboratory and other time oriented data both in research and patient care functions.

At the same time, at the linear accelerator, under an AEC major computer procurement cycle, SLAC was scheduled to receive two 370/168s during the fall of 1973 to augment the 360/91. In the procurement negotiations, SLAC had taken the position that computing would be central to effective physics. The required facility was slated to collect data directly from experiments, spool it directly into digital form, and allow selected interactive analysis by the experimenter of the data in sufficiently short time to modify the experiment while in progress. For back up and reliability, SLAC ordered a dual system. The computer system had to be up whenever the accelerator was up. Otherwise, SLAC priorities seemed entirely like those of the medical school. Also, the SLAC facility was moving out of the exclusive domain of physics. At the request of the director, various administrative functions for SLAC, including payroll and inventory control, were moved to the 360/91, so that many features of production schedules were being imposed on the 360/91.

On the university administration 370/145, the impact of Project INFO development was being felt. The on-line administrative system, OASIS, was being introduced for student services and alumni records (See Chapter 11). The direction in university administrative computing was toward reliable computing, with production schedules, but in a data base management environment and with on-line services.

In the campus facility, the general purpose academic computer, the trend was in much the same direction. A research program on information retrieval under the direction of Professor Edwin B. Parker has led to a product called SPIRES which had been used as a basis for Stanford library automation. In 1972, project BALLOTS was undertaken to develop an on-line, machine based library system using the MARC file and also handling original cataloging in the university library. The assumption was made that if the computer was down, the library catalog department would be down.

The direction of motion in all five computing facilities was much the same. The more conservative management requirements that one would expect in business computing were evidenced in the academic centers, and the administrative facilities were moving towards on-line, interactive computing.

When these facts were considered, the university administration decided to take advantage of Grosch's law, or, "It's cheaper by the dozen." For twice the investment, one can get four times the computing. If an organization needs more than twice the computing, the law seems still to be valid.

COMPUTING SERVICES 1973

In April of 1973, the university established the Stanford Center for Information Processing, or SCIP, to manage all five university service facilities, under the directorship of Charles Dickens, former director of the SLAC facility, with Mike Roberts, who continues as director of the university administrative facility, as Deputy Director. SCIP has moved in the direction of establishing two rather than five geographic locations for computing. One is located at SLAC, which is some miles removed from the campus, and one at the central campus, interestingly enough in the neighborhood of the university library. All computer operations have been placed under unified management. Soon there will be only three installations. The model 50 has been returned. A 370/158 has been established next to the 370/145 and the 370/135 will soon be released from the hospital. The entire medical center will then be using remote job entry and terminal oriented computing. The ACME system is installed on the 370/158. There is a single management for systems programming from all five facilities with a charter to use the temporarily excess hardware resources at SLAC for system development to develop a common environment (interactive services, data base management, real time services, and operating system) for the university as a whole.

Of course, there was before and there is now more to computing at Stanford than these services. For example, in the Graduate School of Business, an H-P 2000 runs BASIC for the MBA students with one operator and a faculty director. The computer is left running unattended nights and weekends, and is very cost effective. In the consolidation, Stanford University has brought together general purpose service computers while continuing to allow unique services. Between 40 and 50 minicomputers, particularly in laboratory support and real time service, will continue to represent the second law of economics in computing. Professor Cox of Washington University has asserted that if one knows what computing one wants, it can be had wholesale. Grosch's law and Cox's law are two primary forces at work at Stanford. The trick will be to get our computing cheaper by the dozen, but also to get it wholesale.

Chapter 2

Management Information Systems

by Donald C. Carroll
The University of Pennsylvania

Eight years ago as an academic concerned with management information systems, I could have elegantly demonstrated, practically proved mathematically, the possibility of generalized MIS for university administration. Four years ago, as general manager of a systems development firm, I would have been delighted to present any of you with a proposal for a generalizable management information system but probably not at fixed price. Now as an academic administrator, I am considerably less optimistic but not totally without hope. It is difficult to say whether it is age, fiscal responsibility, or conscience that has made a coward of me.

Today, I would like to outline briefly the purpose and nature of academic MIS's (I will neglect supporting operations), describe our financial management system at the University of Pennsylvania, outline some of the characteristics of data processing systems drawn from this example, and offer some rather dreary conclusions about the transferability of such systems among universities.

THE PURPOSE AND NATURE OF ACADEMIC MIS'S

Let us assume that the objective of a university is the creation and dissemination of knowledge. Recognizing the current economic situation, that objective might be better stated as the creation and dissemination of knowledge within fiscal constraints. The objective of an information system for managing a university ought therefore to be to provide for efficient allocation of resources toward this end. If the MIS is computer based, data bases and program aids for decision-making must be provided for planning and for control of university operations. What kind of data bases are appropriate particularly for the academic side of a university? There should be an accounting data base that comprehends budgeting as well as control, personnel data bases dealing with both faculty and staff, an admissions data base, a registrar data base, a development data base to deal with alumni and friends of the institution, undoubtedly a research and services data base, and perhaps others. Programs for retrieval and reduction are also needed whether for routine reporting or *ad hoc* retrieval. Extrapolation programs that help to look into the future and simulation programs to deal with various "what would happen if" questions that the managers must ask are also needed. Finally, one needs programs to update and maintain these data bases.

The structure of management information systems must reflect the hierarchy of university organization. Normally, the organization chart for a university begins with a top management typically consisting of a president and provost, continues with division into schools or colleges, is further divided into departments, and ends with the combination of larger departments into groups. In the allocation of resources, however, other slightly different entities within universities such as independent research centers are important within schools, departments, and groups. Sometimes, these are other academic organizations such as graduate groups that may include faculty from several departments not otherwise related organizationally. Many universities and schools now have matrix organizations, with research and teaching program structures cutting across department and schools.

One can perceive differences between state universities and private universities. In state universities, the flow of funds is from top to bottom in a very real sense; the allocation decisions are highly centralized. In the private institution, normally less centralized, the model is generally baronial and funds flow in different directions. While some baronies are capable of contributing to a central fund, other baronies only can exist by drawing from the central fund. (In some private universities, but very few, funds flow from bottom to top.)

Recently it has become fashionable, especially in private universities, to form "responsibility centers" in order to decentralize allocation decisions explicitly. A description of such a system undertaken at the University of Pennsylvania illustrates some of the problems encountered in attempting to generalize computer based information systems for universities.

RESPONSIBILITY CENTER BUDGETING – ONE APPROACH TO MANAGEMENT

For a responsibility center manager, for example a dean of a school, the fundamental set of information is the responsibility center report which is really a profit and loss statement. It consists of various income items: tuition less financial aid; fees; endowment income; gifts; grants and research contracts; income from sales and services (very important in medical schools that have hospitals); and an item called interest income. On the expense side are salaries and wages (whether administrative, academic, secretarial, or part time employees); fringe benefits; various current expenses; equipment (especially important in the physical sciences); computers; and interest expense. At Penn centers are asked to adjust expenses to be less than income and, in fact in some case, substantially less than income. It is hoped that the difference between income and expense (the "target") is sufficient to cover indirect expenses: libraries, building operation and maintenance, the President's office and the like. Parenthetically, centers vary greatly in their capability and attitude toward contributing to indirect expenses. In the budgetary cycle, the central budgetary staff works out targets which are passed to responsibility centers. These are expressed as a difference between direct income and direct expense as a first step in a process of budget negotiation. The responsibility center manager is expected to respond in one of several ways. He can accept his target and produce a supporting statement that shows how he will meet the target, or he can say, "I cannot make that target," and negotiation follows. At some point, approval is obtained and a budget is established for the next fiscal year.

During the year, managers control to a "flexible budget." That is, if more tuition is earned, more endowment is obtained, or more research contracts brought in, it is possible to adjust the expense side of the budget, typically, by acquiring research assistants, visiting faculty, and teaching assistants. At the end of the year, according to the rules of the game at Penn, if the target is exceeded by the responsibility center, that additional money accrues to that center. It is deposited in a "bank" and draws interest which is shown as interest income on the profit-and-loss statement. If the center fails to produce to the target, a deficit is carried to the bank and interest is charged to the center yielding interest expense.

The responsibility centers system does not stop at the level of schools but is to be carried down to the level of departments or research centers which may, in turn, wish to set up subsidiary responsibility centers. It is thus a fiscal control system, because the allocation of resources is broken down to subsidiary levels, an academic administrator does not have to run to the provost to trade money from the academic salary account to the secretarial account, for example.

IMPLICATIONS FOR DATA PROCESSING

Technically, the various line items on the profit and loss statement are derived from many diverse data bases. For example, the line item, "tuition," is derived by a very complicated process drawing on the registrar's data base and the student aid data base, classified by department or group within schools. This type of system which is typical for many private universities has certain implications for data processing. If subsidiary responsibility centers are created, each transaction or each entry into the data processing system has to be coded at a very low level. For example, a tuition credit must be recorded for each particular group and department within a school. Multiple activities of faculty have to be carefully tracked and accounted for. An individual faculty member may have affiliations with three groups, for example. One must have an extremely adriot bean counting operation to code a transaction before even considering processing the transaction.

Let me attempt some generalizations on the characteristics of systems that can comprehend the complexity of a management scheme of this type. Most of the time, time is not of the essence. "Real time" in an academic organization probably can be defined with a response time of several weeks, not milliseconds. However, the complexity does place a tremendous premium on *flexible* retrieval from the system. One would want to search for information on many, many dimensions: the student; the faculty; the research center; the department; the research project; the research sponsor; the program; and probably by type of activity. Possible inquiries are literally endless, and it is difficult, if not impossible, to cover the needs of the academic manager with routine periodic reports. Flexible retrieval requires realtively straightforward report generation or, in the more elegant situation, on-line retrieval.

CHARACTERISTICS OF ACADEMIC MANAGERS

One must also take into consideration the characteristics of the managers of the academic community. The most notable characteristic of academic managers is idosyncracy. Among academic departmental chairmen, there is a diversity that would be astounding in a commercial organization. It would be nearly impossible to change this or to regiment the way that academics manage academics. A computer-based MIS must provide for such a variety of managerial style. One must also understand that academic managers are transient. The average half-life of a department chairman and deans is now on the order of two to three years. Hence, a system *must* be easy to understand and operate.

CONCLUSIONS

Computer-based MIS systems designed to serve these departmental chairmen must deal with large and very carefully maintained data bases with requirements for an atomic level of data somewhat more detailed than the typical commercial operation requires. The system also will have to produce few routine reports relative to the *ad hoc* reports that are required in the process of management, and it must consequently have a variety of retrieval and reduction programs.

The system called "CIS" by the IBM Corporation was an elegant design that met many of these specifications. It was probably not a commercial success because it cost too much. I fear that it is the appropriate example to cite, that is, university MIS's, if done right, cost too much. Most of us will have to settle for less. Thus, I predict that each university will want to make its own compromises and respond to factors that are most important to it, and factors important to one university are unlikely to be most important to another. Where teaching is the fundamental activity and money is not particularly tight, one kind of compromise will be made. Where research is fundamental and budgets are tight, a totally different type of system will result from the compromise between requirements and money made there. Because of these different compromises I suspect "exportability" of academic MIS's will be difficult and limited. I hope I am wrong, but I do believe that we should recognize the inherent problems in our pursuit of management information systems that "can travel."

Chapter 3

Computers in Instruction

by Charles H. Warlick
The University of Texas at Austin

The University of Texas is probably better known in terms of football than it is in terms of instructional computing. However, there is a great deal of instructional computing at the University of Texas and it has been, to a great extent, incorporated into the curriculum at the university.

For example, computer-assisted instruction is a real on-going practice with hundreds of students in chemistry and geology spending thousands of hours at console terminals interacting with instructional modules. To provide this service, the systems staff at the University developed a conversational language for instructional computing, which is a superset of Coursewriter II and FORTRAN IV, along with all of the supporting CAI management software.

With a new computer-augmented lecture technique, instructors are using the interactive system with large rear-screen projectors before large classes. This permits instructors to carry out complex calculations before the class, instead of leaving the difficult problems as class exercises.

The aggregate of instructional computing activities can be seen in the statistics for undergraduate and graduate classroom use compared to the total computing activities last year. Forty-five different academic departments used the interactive computer system in 1972.

Academic computing has been centralized at the University of Texas-Austin for many years. A seven-year-old Control Data 6600 and a younger 6400 are connected through a large core memory and through a large mass storage system. The computers operate independently but share a common file system. They are used exclusively for academic purposes with all data processing being handled in a separate business office computer. The 6400 serves interactive users through 128 timesharing ports, and the 6600 handles the batch load through 16 remote batch entry terminals. There are also other computers serving special academic needs on the campus, but nearly all are interconnected with the 6600 forming a campus computer network. Like most institutions with large computer facilities, Texas has used its computers primarily for support of large research projects, particularly in the natural sciences. However, in recent years there has been a substantial up-swing in the use of computers in instruction.

Figure 3.1 Computing Activities 1972-73

900,000	Programs
7,300	Hours of Computer Time
18,000	Undergraduate Students
600	Classes
330,000	Programs
32,000	Hours on Timesharing Terminals
1,200	Graduate Students
180,000	Programs
2,000	Hours of Computer Time

Texas served as the host or supplier institution for nine other colleges and universities in one of the National Science Foundation regional computer networks during the three academic years 1969 to 1972. Today 23 institutions participate in this educational network, which is operating on a self-sustaining basis.

The success of the Southwest Regional Network was similar to that of other regional networks: terminals and communication lines were installed and checked out, and interactive and remote batch services were brought to a stable, reliable state. Unfortunately, however, most of the network projects were too short to do much more than get underway, and even then a number of them faltered and disappeared.

The really important questions relating to the effectiveness of the academic computing in the instructional process went unanswered for the most part. Therefore, the NSF launched a separate project addressed to these questions with the EDUCOM study of *Factors Inhibiting the Use of Computers in Instruction* which was directed by Ernest Anastasio.

In the final report, Mr. Anastasio stated that the major question to be answered was "how can *evidence* of effectiveness be provided?" The following circular sequence of answers evolved: To provide evidence of effectiveness, one must conduct a convincing high quality demonstration, but this requires good computer materials, and good materials can only be prepared by people who know theories and methods of instruction. To get good people, one needs professional recognition and economic incentives, which in turn require evidence of the value of the pursuit and a formal production-distribution system. These require a market. To get a market one needs a demonstration of effectiveness, and the loop is closed. Hopefully an infusion of funds somewhere in the loop would help break the cycle.

The National Science Foundation has joined the University of Texas in an attempt to break this cycle by producing quality computer materials and demonstrating these materials. The Computer-Based Education Project has been established, which is a five-year curriculum development effort with four primary goals: (1) to identify the common concepts among disciplines; (2) to develop evaluation schema; (3) to develop transferability criteria; and (4) to develop an implementation model. Directed by Dr. J. J. Allan and Dr. J. J. Lagowski, Project C-BE is applying current technology and currently available devices in concert with sound pedagogical practice.

Ultimately Project C-BE will involve 75 professors and over 4,000 students in 44 different curriculum development and demonstration projects. The project is entering its third year and its impact is readily apparent. For example, during the Fall 1973 semester 25 of the sub-projects will test computer-based instructional modules with 1,200 students averaging 2,000 console hours per week. A wide spectrum of disciplines is represented including physics, chemistry, psychology, engineering, statistics, biometrics, linguistics and home economics. Without any question Project C-BE is helping to bring about changes in the educational policies of the university, including a change in the university attitude toward the allocation of computer resources.

The National Science Foundation has also launched another project to break the Anastasio cycle by attacking the problem of the formal distribution system. The NSF brought together five of the successful regional computer networks, including Texas, to undertake a study of transportability and dissemination of computer related curriculum materials in project CONDUIT.

The five computer networks comprising CONDUIT include those centered at Dartmouth College, the University of Iowa, the North Carolina Educational Computing Service, Oregon State University, and the University of Texas at Austin. Each institution offers a unique set of resources; no two have the same type of computer system, and together they provide instructional computing services to 275,000 students at 100 institutions of higher education. Representing almost 5% of the institutions of higher education the CONDUIT networks provide an ideal experimental base for a study of the transportability and dissemination of computer-based teaching materials.

CONDUIT OBJECTIVES

The *visionary goal* of CONDUIT is to use the computer to improve undergraduate education in a cost-effective manner by the exchange of computer related curriculum materials. This project was concerned with the belief that exchange can significantly multiply the benefit of expenditures on curriculum development. However, the project directors realized that an attack on this problem required a *study* of the movement and dissemination of curriculum materials which use computers in undergraduate education. Movement and dissemination then became the *operational goal*. The curriculum materials being studied run the gamut from simple computing exercises, to instructional modules, fully integrated curricula, computer-oriented texts and other adjunct materials. In planning the research program to study the movement of curriculum materials, the text book publishing industry was taken as an excellent model. Unfortunately, though, there have been numerous failures in attempts to use this industry for the distribution of computer-related materials.

CONDUIT's primary objective was to look for the requisites for achieving transferability by looking at different methods of transfer. This, in turn, broke down into 10 sub-objectives:

1. Creating dissemination strategies which differ in the manner in which they perform their various functions.
2. Obtaining quantitative measures of "success" of dissemination.
3. Determining subjective aspects of computer based material dessmination such as acceptance and attitudes.
4. Determining guidelines for technical transport.
5. Establishing a small, high-quality reservoir of materials based on experimentation with disciplinary review and technical verification of the materials.
6. Publicizing the availability of materials.
7. Obtaining cost-effectiveness data.
8. Determining the irreducible minimum of procedures for dissemination.

9. Providing insight into the human interrelationships that must necessarily accompany distribution activities on a national scale.
10. Determining which CONDUIT services could be made wholly or in part self-supporting and what the long-range role of CONDUIT should be in servicing the dissemination of computer-related curriculum materials.

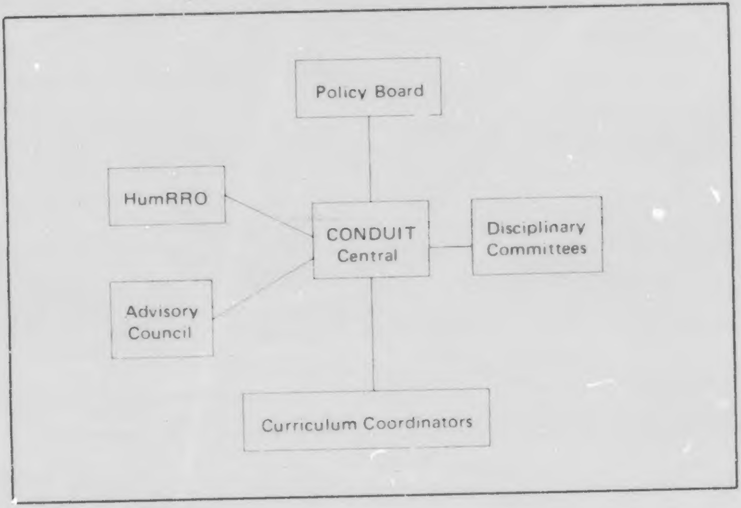
The intent of CONDUIT was to accomplish as many of the objectives as possible during the study phase of the project.

The determination as to whether these objectives are met is a responsibility which should not be done internally but rather should be vested in a body outside of the project staff conducting the experiment. The Human Resources Research Organization (HumRRO) which has undertaken the evaluation function became an important contributor to the design of the vehicles through which data could be collected for evaluation.

CONDUIT ORGANIZATION

CONDUIT has developed a formal structure to accomplish its goals. (See Figure 3.2)

Figure 3.2 CONDUIT Organization Chart



The *Policy Board* consists of the directors of the five computer centers that make up the CONDUIT consortium: Larry C. Hunter of Oregon State, Thomas E. Kurtz of Dartmouth, Louis T. Parker of North Carolina ECS, Gerard P. Weeg of Iowa, and Charles H. Warlick of Texas. After a two-year period of rotating the chairmanship, Professor Weeg has been named the "permanent" chairman of the Policy Board. As its name implies, the Policy Board is the ultimate authority for all actions in CONDUIT. As such it sets or approves the long-range goals and activities of the organization.

CONDUIT Central is the executive branch of CONDUIT charged with the direction of the activities of CONDUIT, the two-way liaison with all constituencies of CONDUIT, the creation of good, new directions and activities for CONDUIT, and the liaison with the evaluation agency, HumRRO. Dr. Jim Johnson of the University of Iowa is the current Director of CONDUIT and heads the CONDUIT central organization. During the first phase Dr. Ronald Blum served as CONDUIT Director.

The *Curriculum Coordinators* located at each of the five nodes of CONDUIT are changed with implementing all local aspects of CONDUIT activities.

These three components were envisioned in the original design of the CONDUIT organization. As the project progressed, it was recognized attention should be concentrated on a small number of disciplines, and that help was needed from the leading proponents of instructional computing in those disciplines. Seven disciplines were selected for which computer-based curriculum materials were known to exist: Biology, Business, Chemistry, Economics, Mathematics, Physics, and the Social Sciences.

Committees were then set up for each of the disciplines consisting of four persons with experience in and a commitment to the use of computers in the undergraduate teaching of the discipline. The *disciplinary committee* members are not necessarily from CONDUIT schools. Each committee has elected its own chairman, and the seven chairmen and Dr. Karl Zinn of the University of Michigan constitute the *Advisory Council.* This council provides the Board and CONDUIT Central with a panel which reviews plans and activities and suggests new directions for investigation.

In fact, the establishment of the disciplinary committees is one of the most important things the project has done. The committees have worked wonders in locating existing computer-based curriculum materials in their disciplines. They have reviewed them and passed on to CONDUIT recommendations of materials worthy of consideration in a program of testing and evaluation.

The organization diagram does not show the indirect but important contributions of the NSF project Monitor, Dr. Andrew Molnar.

CONDUIT ACTIVITIES

Specific activities carried out by CONDUIT thus far include regional workshops, national classroom tests, hypothesis tests, video-tape seminars, and independent transfer of computer based curriculum materials.

Each of these has the dual role of providing information to identify critical factors in the movement of computer-based curriculum materials, and helping to determine the relationship between dissemination strategies and user adoption.

Generally, the classroom tests have been preceded by *national workshops* in which faculty from schools in the CONDUIT networks are instructed in the pedagogical design and use of curriculum materials passed on by the disciplinary committees. The CONDUIT nodes have assumed responsibility for activities in one or two disciplines and the workshops for the different disciplines have been held at these sites. For example, the chemistry materials are handled out of Texas. The attendees at the workshops, however, come from all of the CONDUIT networks.

The workshop faculties have usually included the authors of the computer-based curriculum materials which will be tested by the attendees back in their respective classrooms. The authors have discussed their approaches in the development of the materials, techniques of their use and experiences with the materials.

The materials presented at the workshops have been installed at all of the CONDUIT networks, and the workshop attendees have now conducted many classroom tests of the materials. This *national classroom test* effort has been carefully monitored by a series of 18 data collection instruments ranging from a programmer transport log to a student evaluation of materials. [illegible] collection and analysis of data are being performed by CONDUIT with the results being evaluated and reported by HumRRO.

The national workshops have been followed-up with regional workshops conducted by participants in the national workshops. This secondary effort has been very important in spreading the classroom tests throughout each region.

Working with HumRRO CONDUIT has also developed a set of *hypotheses* regarding the transportation of curriculum materials. To test the validity of the hypotheses data generated from the classroom tests is used and specific activity toward testing hypotheses has also been undertaken. Three of the ten hypotheses are:

- A market for educational computer usage and program exchange does, in fact, exist.
- Workshops are necessary to provide a unique combination of motivational and informational factors for the user.

- The greater the physical and psychological distance of the potential user from a CONDUIT network, the less likely the chance of adoption of the educational technology.

Specific activities undertaken to test hypotheses include design of alternate dissemination systems for materials in biology and social science. These materials were put in self-demonstrable form to compare this form of introduction to workshops. Notification to make faculty aware of the self-demonstrable materials included brochure mailings and telephone contact. As a control, workshops were held in two different regions.

Some independent transfer has occurred involving the movement of materials that were not part of the formal CONDUIT classroom tests. This activity has been initiated by CONDUIT personnel who have made use of interpersonal relationships cemented by the CONDUIT study.

CONDUIT PRODUCTS

The entire evaluation effort of CONDUIT is directed toward producing several *final products* that will be useful for developers, for persons involved in transporting materials, and for users of transported curricula. User products are designed to provide the academic community with information that will be valuable in facilitating curriculum innovation using computer based materials.

Documentation guidelines will be prepared to provide a systematic set of standards for documentation of computer-oriented materials.

Selection guidelines will establish standards to guide selection of computer-based materials to be moved from one site to another.

Technical verification guidelines will define procedures for determining the correctness of a computer-oriented package after it has been imported and installed in a computer facility. Such procedures will serve to ease the process of testing and detecting errors that are either the result of technical mistakes in translation or the result of machine dependent characteristics of program operation.

Technical transport guidelines will contain a list of the common problems encountered in movement of computer-oriented materials.

An important by-product of the CONDUIT experiment will be the libraries of tested and validated curriculum materials installed and in use at the CONDUIT networks and available to all who wish to acquire and use them. These will be complete packages of programs backed up with all appropriate textual materials.

An on-line catalog of available materials has already been established on the Dartmouth computer for use, at present, by the CONDUIT participants.

ACCOMPLISHMENTS TO DATE

In little over a year and a half CONDUIT has defined the major factors involved in the movement of computer related curriculum materials. Hypotheses related to these factors have been established and activities to test the hypotheses have been initiated. CONDUIT has collected data on activities and published preliminary findings and guidelines suggested by the data. A major activity has been the selection, dissemination, and local installation of 87 program modules. Two integrated courses, where computer work is an integral part of the entire course have been transported, and nine data bases in seven disciplines have been installed for use by over sixty classroom instructors. Some 7600 students have been exposed to CONDUIT materials, and 500 questionnaires from students have indicated positive attitudes on their use.

The results of this experiment together with other CONDUIT activities will provide valuable information about the dissemination process heretofore unavailable to the academic community. More significantly it will suggest cost-effective methods to overcome psychological and technical barriers to the movement and use of computer-based instructional materials.

While the five networks comprising CONDUIT share an interest in the experimental results of the CONDUIT study, their involvement is more basic. Each is convinced that higher education can be significantly improved through the use of computing, and each is determined to spearhead that improvement. They have devoted personnel, resources and creative effort to the CONDUIT experiment that could have been directed toward other projects. They are determined that during the third year, CONDUIT will product the products mentioned and a definitive report on the dissemination process that will give effective direction to future transportation of computer related curriculum materials. They also anticipate that the CONDUIT project will expand as additional schools enter the experiment.

TEXAS OUTLOOK

At Texas the attention of the academic administration has turned to instructional computing in a very real sense. For example, the Provost has recently established a committee on computer-based instruction to advise him and the deans on directions to be taken on the use of computers in instruction. But the real evidence is allocation by the Board of Regents of one million dollars for the further development of instructional computing during the next twelve months. The University-wide Faculty Computer Committee requested the funds and is responsible for their allocation. Approximately one-quarter of the amount will go into console terminals in new teaching centers to be established in the various colleges of the University of Texas-Austin. About $500,000 will go into the expansion of the interactive service, with particular attention to the support of students interacting with instructional modules. The balance will go into remote batch computer terminals and equipment for special teaching projects.

The University of Texas is in a unique position today with regard to instructional computing. This uniqueness derives from the University's large capital investment in academic computing, from the administration's drive to incorporate this technology into the curriculum, and from the direct thrusts of Project C-BE and CONDUIT, which will further incorporate computers into the instructional program.

Texans do not undertake things in a small way. The University expects large successes, and expects to share them for the general benefit of higher education.

Chapter 4

Computing in Research

by Robert L. Ashenhurst
The University of Chicago

In developing my notes for this address, I decided that it would be more appropriate to include the word "computing" in the title, instead of "computers" as appears in the program. In this way we are led to concentrate on fundamentals rather than equipment. It is a never-ending source of surprise to me that any discussion which begins by addressing what research objectives should be supported by computing still tends to end up with a discussion of the latest hardware acquisitions of this-or-that computing center.

It is appropriate to characterize the decade of the 1960's as that in which computing in research came into its own. Earlier, in a time which might be characterized as the Dark Ages, automatic computing facilities were available to only a few researchers. At that time the potential of the computer as a research tool, although obvious to some, was certainly not widely appreciated. By the end of the decade, with the flowering of the Renaissance, everybody was carrying around printouts and sitting at terminals. The essential difference, although the expansion of physical facilities was substantial, was the rise of the notion that computing or access to computing is an integral and indispensable part of research for a broad variety of disciplines. The growth of physical facilities was inextricably linked with the development of this notion, and the pattern of that growth has given us a legacy for the future.

But the situation is again changing, as suggested recently by Martin Greenberger: "The centralized operation that tried to be all things to its broad spectrum of users within the institution is giving way to extra-institutional approaches to providing and receiving information and computing services."[1] It is becoming fashionable to declare the university computing centers that grew up in the environment of the 1960's to be outmoded, and to recommend that the resources supporting them be turned to more up-to-date schemes like networking and hierarchical computing.

It remains to be seen if in the decade of the 1970's there will be a Reformation and a Counter-Reformation, or whether all that can be skipped in getting on with the Age of Enlightenment. The problem is to effect a transition without resorting to violent means toward the protectors of the old order. At the same time it must be recognized that the apostles of the new way can lead us into a new set of systematic rigidities, unless their doctrine is viewed in the context of the realities of the fundamental nature and purpose of research computing.

Computing centers that grew up on university campuses sometimes started out with a certain confusion as to mission, characterized by a different understanding on the part of those who would use the system for research from that of those whose budgets had to support the center. Eventually a pattern emerged whereby the computing center was a recognized source of a particular set of services to the research community. These were well characterized as "computer systems services" or "computing services" rather than "applications services" specific to research in a discipline. Computing services involved keeping the hardware and the basic operating software up and running, and providing high quality, fast turnaround batch processing and later interactive access, using a variety of programming languages. The provision of these services became professionalized in the customer type relationship of the researcher to the computing center.

In the ideal case, this was an appropriate structure. We can think of the researcher as interacting formally with the center, expecting the computing services to be available reliably and reasonably. Although the researcher grew quite dependent on those services, he still considered the content of the programs and the data files that he used to be basically his own responsibility. He interacted informally with his students and colleagues to develop and exchange ideas and programs. If he wanted to obtain programs from elsewhere, there was a problem getting them to run at his home center. Sometimes it was necessary to go to another installation, preferably in California, to run someone else's programs in their natural environment.

In the situation described, both the dedicated group of people who supplied computing services (managers, systems programmers and operators) and the dedicated group of researchers (research investigators, co-investigators, research associates, and students) were doing something in which they believed. However, matters in practice did not always measure up to the ideal. Disagreements over what constituted good computing service often arose. Computing center managers were sometimes left defending a not very defensible position due to idosyncracies of their staff over which they had no control, and researchers lost no time in incorporating these idosyncracies into their store of anecdotes about the frustrations of using centralized computing facilities. In addition, more systematic pressures acted to increase the gap between what the user thought he wanted and what the center was wont to provide. The management ground rules under which the center operated usually placed a premium on catering to a stabilized set of demands. Thus more routine uses, often administrative data processing, were emphasized at the expense of service to researchers who tended to have variable demand and financial support for computer use. Management ground rules also acted to inhibit innovation unless a very clear demand for a new service existed.

Even with these limitations, functioning in this mode was more satisfactory than when the roles of the two dedicated groups were confused. Often attempts by a generalized computing center to supply services that intimately tied into research applications tended to be even harder to maintain and manage effectively to the satisfaction of users. Such service programming for researchers was not terribly welcome. At the same time, attempts by academic departments to run individual computing facilities, although occasionally successful, usually did not work. Because a computer standing in the corner in a department is not evident to the world, these failures were not as apparent as deficiencies in a computing center used by the whole university.

Today, the ambivalence that characterizes these conflicting trends in the development of computational support of research is aggravated by several new features which further complicate the picture, but which, properly brought together, are capable of leading to enhanced capability over a broader range of research needs. These features, naturally enough, represent trends in opposing directions. Two paramount trends are those toward use of minicomputers and toward nationwide discipline-oriented computing packages and databases.

The most immediately evident new factor is the minicomputer. Minicomputers are frequently used for realtime data acquisition, in which case they become part of the experimental instrumentation and cannot be denied to the research investigator in the name of centralized computing. A basic minisystem costs $5,000-$10,000, but more equipment is often applied for when it becomes apparent that the basic system does not provide all the services that the investigator needs.

With such a minisystem, the notion of a departmental computing service becomes viable. A research investigator who has his own computer may say, "My computer is available. I'll let all of my group use it as they want. All my graduate students will learn how to use it. It will be there when they want it on a first come, first served basis." Most computing center directors find such an arrangement difficult to counter, especially when the person who wants to acquire the computer has the funds to do it. The researcher is thus able to go back to the style he knows so well, that of the informal facility. Of course there is a problem if the original investigator leaves the institution. Then what could be done easily yesterday cannot, for reasons not easily discernible, be done today. That, however, does not seem particularly strange of people who like to do computing systems research as an ancillary activity to their disciplinary research.

Additional factors, however, have produced a countertrend to the minicomputer revolution: the availability of computing packages applicable to a particular discipline; and the development and availability of discipline-oriented databases. Researchers have put together systems of programs and aggregations of data that are of use to a wide variety of other workers in their discipline, and have committed themselves to maintaining and disseminating these materials. Note that here a third type of dedicated group has appeared, a group of people in a discipline who in fact are spending their time on making research materials available rather than continuing discipline-oriented research. Although problems do develop as to professional motivation, such packages and databases are being brought forth in increasing numbers.

The emergence of discipline-oriented resources leads to the conclusion that large computer networks are the way of the future. Because users can tap into a remote system as if it were local, it is possible to access packages and databases that physically reside elsewhere and to avoid the ever-present difficulties of transporting them to one's own installation. The technical feasibility of such networks has been demonstrated by ARPANET and other prototype developments.

As emphasized in the recent EDUCOM-NSF sponsored seminars on computer networking[2], the problems of making networking a reality are "political, organizational and economic." As disciplinary centers such as the NCAR facility for geophysical and atmospheric sciences are established, their existence will certainly be a political and economic factor that will facilitate the evolution toward a network-type economy, despite the substantial obstacles that still exist. It is also customary to consider organizational problems as being formidable. When considering networking, research users often compare network use to local computing center use saying, "Well, we have a certain amount of trouble using our own computing center. How are we going to tap into a terminal communicating with a remote computing center? We don't know the people there, we don't even know their phone number sometimes, and we wonder where they are on this particular morning." The solution to this type of problem, while not primarily technological, can nevertheless be approached through technology.

There naturally emerges the notion of a hierarchical computing system where one has access to an intermediate-level local computing facility, backed up by a network of higher-level computing services. It is frivolous to give a researcher access to a myriad of remote systems without mediation. He needs services that are more appropriate for his interests. There has been a certain amount of talk about the wholesaler and the retailer of computing services. Many of the organizational difficulties of providing computing services through networks can be dispelled if the retailing function is organized as a dedicated group with the objective of making those services available to local researchers.

This opens up the prospect of a variety of new system configurations, both hardware and software, for local nodes of networks serving particular communities of users. Under the rubric of "networking" and/or "hierarchical computing," many projects have come into being in the past two or three years to exploit the possibilities. Some of these are supported by government grants, others represent attempts by campus and other computing centers to broaden the class of services they provide. Unfortunately, but typically, the terminology of characterizing such configurations has not yet matured to where the same words mean even roughly the same thing to all parties. Programmatic efforts by the National Science Foundation and the National Bureau of Standards and others, including EDUCOM, should help to get the situation more sorted out in the near future.

A research group at the University of Chicago is attempting to implement such a function for a system being developed as one of several projects supported by the National Science Foundation. A detailed discussion of the system is given in the final report of the EDUCOM-NSF sponsored Seminars. The Minicomputer Interfacing Support System (MISS) is conceived as a specialized facility for serving researchers who need minicomputers in their experimental investigations, although it may support minicomputers used in other ways. Often a researcher with a minimal minicomputer system will find that he has temporary needs for all sorts of additional facilities which can lead to unanticipated expense if he has to acquire them for himself. A hierarchical system like MISS serves such minicomputers that investigators acquire themselves and plug into the system. The system supports a variety of minicomputer types, and permits the investigator who has a minicomputer to have it online or to use it standalone as his needs dictate. The system is designed to serve investigators who will use the system when they need its services, and not otherwise.

At the highest level of this system there is a large central computing facility. At the present this is the University Computation Center, which has recently upgraded to the IBM 370/168, but it is hoped that eventually this will be expanded to a network. Access to the network is through an intermediate level, which is a specialized aggregate of hardware and software and an associated systems staff. This staff is thus the link between the researcher who has the occasional use for a higher facility but does not want to be bothered with a lot of intricacies, and the large centralized facility or network, which is supplying the computing services. Technologically, one could have all levels above the minicomputer be served by one large centralized facility, but a very important part of the MISS design is the intermediate level in the three-level system, because of the need to facilitate the interaction of minicomputer systems with higher level computing services.

Thus hierarchical computing is more than accessing various levels of computing power. Essentially it is provision for maintaining specific services at the various levels through specific organizational features as well as communication lines. Hierarchical computing organizes, or helps to organize, the groups of people so that the service purveyors can concentrate on purveying services and the researchers can concentrate on research. In this way best use can be made of the new technological developments that have emerged in the last few years, to permit their claimed potential to be realized in benefits to users.

The University of Chicago MISS project embodies only one way of organizing a hierarchical computing system, and such systems represent only one way of providing new types of computing services in support of research. It is hoped, however, that the present discussion has indicated the need for reevaluation of the traditional attitude of researchers toward computation facilities as viewed by many, and for keeping an open and inquiring mind concerning the possible scope of "research computing" in the future.

REFERENCES

1. "Computing in Transition" *Science*, 28 September 1973
2. Proceedings to appear as *Networks for Research and Education*, Greenberger, M. et al (M.I.T. Press, Cambridge, Mass.) 1973.

PART II

ADMINISTRATION
OF
COMPUTER
RESOURCES

35/36

Chapter 5

Administration of Computer Resources: An Overview

by James Poage
Princeton University

One theme which links together all the talks on Computers in Research is the fact that computing for higher education is in a period of change.

Basically, there were four ways that speakers looked at change. First, some described what has been done with resource sharing. Another class of talks reviewed plans for resource sharing or improvement of services. The third class included statements like, "If you're going to do it, consider these following things." People came up with a list of suggestions, particularly for college and university administrators of things that ought to be considered in the decision making processes. And finally, a fourth class which focused on statewide plans came out with the warning, "If you don't do it first, somebody else is going to do it for you."

John Skelton from the University of Denver has been looking at networks. He is in the process of publishing a monograph, "Games Universities Play." His remarks may have sounded lighthearted, but he is deadly serious (see Chapter 6 for a full version of his paper). Games such as, "The Balance of Payments" game, the "High Society" game, the "Welfare Versus the Proud/Poor" game and the "Godfather" game, illustrate the ways in which faculty and administrators at universities often interact. If one is going to get into any sort of change situation, such as moving into a resource sharing position, it is best to know how decisions are being made and where they are being made. If one doesn't know this, decisions are not going to be made the way one wants them to be made.

Dennis Fife, from the National Bureau of Standards, reported on an investigation of network management and how it should change as resource sharing moves through various phases. The study focused particularly on the management of change. Dr. Fife isolated five phases or steps that universities appear to go through as resource sharing networks are developed. The first and most basic step is the mutual service arrangement where two or more centers act as each other's facilities. This is usually done on a straight vendor-customer relation. The next step is a mutual support phase where, possibly through some agreement or contract, one institution guarantees another a certain amount of income in return for a reciprocal guarantee. In a third phase, operations coordination begins. Things like cataloged procedures are standardized. The fourth step is taken when a services alignment is arranged where various groups start to specialize. One institution might specialize in timesharing, another in batch processing, and another in graphics. Finally, institutions reach the top level of cooperation when joint ownership is developed, often quite apart from the usual university structure. The five-step process is interesting because it is typical of many real situations. Rutgers University and Princeton University had a lot of trouble in the State of New Jersey a few years back trying to cooperate. Endless committee meetings were held trying to achieve some degree of resource sharing. Looking back, it is obvious that representatives were trying to establish a monster organization that fell between the fourth and fifth categories of services alignment and joint resource ownership. All this was tried without first having established a basis for resource sharing by passing through the earlier phases. Because the institutions tried to move too fast too soon, it just never worked. Under the current arrangement, the EIS organization falls very low in the hierarchy of resource sharing. The facilities management arrangement between EIS, Princeton and Rutgers is a very simple one but will probably be the basis from which more extensive cooperation can be developed within the state.

Charles Mosmann, who is conducting a study of Academic Computer Planning in the States for EDUCOM, classified networks into four categories. In the empire model, which is based on the Dartmouth experience, there is one large central university computing facility upon which everyone builds. In the alliance model, which is based on the TUCC experience, a cooperative kind of network operates. The common market model is typified by ARPA where colleges on the network compete with one another. Finally, the legislative net is achieved when people are forced into a sharing environment by an agency of the state government. Computing for higher education in the state of New Jersey matches closely the legislative net. These models are most interesting in the context of change. If computing services for higher education are changing in the direction of resource sharing, one really should be aware of the differences that exist among these four kinds of nets. If the differences are not recognized, goals cannot be well spelled out.

In two sessions, various statewide plans were spelled out and critiqued by members of other state organizations. For two of the states involved, these were a ceremonial bloodletting.

For conferees, the sessions were very profitable since they could look at states with fairly well-established resource sharing networks that have gone through a very rapid and drastic change. Illinois for example, and New Jersey have been in business a full year now. Oregon is in the planning stage. Clearly state plans are very, very dependent upon the political environment and on the people in those states. What is structured for one state may be totally unrealistic for any other state for reasons that are not always apparent. For example, in 18 months' time, Illinois consolidated twenty-two data processing centers into five and will further consolidate to bring the total number of centers to two. That is traumatic change! Furthermore, the educational data processing budget has dropped from $13 million to about $10 million in one year. That, too, is a trauma! According to Paul Sire, the situation in New Hampshire is very similar. One might draw a moral from these experiences. "Do it for thyself, or the state will do it for thee."

State plans are so dependent upon the individual state environment that one can't really pick up a plan from one place and put it down in another. However, action must be taken by colleges and universities now to build resource sharing arrangements. If universities don't take the lead in cutting budgets and sharing resources, the state agencies will do it for them.

Chapter 6

Alternatives in Computer Administration

The Games Universities Play

by John E. Skelton
University of Denver

In 1971 members of the Mathematics Department and the Industrial Economics Division of the University of Denver were awarded an N.S.F. grant to study alternative methods of organizing, managing and financing computing at the nation's institutions of higher education.[1] Because of several factors including the economic crunch, increased state involvement, and decreased federal support, the investigators felt that the state of computing at colleges and universities was at a turning point. In computing, the gap between the haves and the have nots was widening. As pointed out by Professor Greenberger in a recent issue of *Science*, "The day of the large general-purpose, single facility computer center may be ending..." Although flourising for a number of years, "...a sharp drop-off in external support and funded usage compounded by some customer relations and credibility problems arrested its development. As deficits began to appear like warts on soaring computer budgets, institutional executives started searching for alternatives to the computer center."[2]

Many alternatives to the computer center including alternative organizations, management and financing either had been tried in the past or were in used in an effort to solve the increasing problems. However, information as to the attractiveness, inner-workings and relative success of these alternatives on a campus was not being disseminated to the higher education community at large.

The University of Denver study was conducted through teams site visitations. A typical team was composed of a computer scientist, an economist and an outside consultant (usually someone recognized to be knowledgeable to institutional computing). The team would typically spend two or three days at an institution interviewing students, faculty, administration and computing center staff. Anonymity was guaranteed to all interviewed and a free exchange of viewpoints was invited. In some cases, where specific information was needed, shorter visits or telephone interviews were used. Governmental units associated with computing on a statewide level were also interviewed. The overall project used an advisory board of recognized computing authorities with extensive institutional computing experience. Sites were selected to be visited on the basis of known or suspected employment of a particular alternative of computing organization, management or financing and with the concurrence of the advisory board.

The final report will be a pocketbook-sized document of approximately 150 pages written to be read by college and university presidents.[3] This report will *not* be written for the computer scientist or the computing center manager and these specialists will probably disagree with much of what is said. The report will summarize what is now going on in computing at institutions of higher education. It will indicate how to recognize the current state and future direction of computing at an institution. It will analyze the various pressures and pressure groups involved, who really determines what is offered, and more importantly not offered, and why. Hopefully this report will give the beleaguered top administrators some feel for what is going on around them and what questions to ask of their own computing specialists. It will not tell the administrator what is the right or wrong decision, for only the institution can determine that. What the report will give is a menu of alternatives and some ammunition. It will give some indications of what is good about an entree and what is bad, what the price is and what the benefits are. The report will close with several scenarios taken from real situations. The names, obviously, will be changed in order to protect the guilty.

At first blush the title "The Games Universities Play" may seem facetious, but is in fact quite serious. There are many similarities between what can be observed going on in institutional computing and the content of the late Dr. Berne's popular book.[4] Colleges and universities have been faced, and are faced today, with various computing alternatives. Computing decisions are not usually based upon technologically oriented arguments but rather are formed as a result of the dynamic human and organizational processes within the institution itself. A game is being played by the parties involved in the decision and the game deals with the interaction of computing alternatives and established institutional decision making processes. The games determine what computing alternatives are and are not considered, why they arose in the first place and how the issue is resolved. The game is played quite seriously, either knowingly or unknowingly, both by administration, faculty and staff who are quite intelligent and by those who are not so intelligent. It is played by those who have the best interests of the institution at heart and those who have their own best interests at heart. The stakes are high and the outcome and total effect of the game at an institution may not be known for quite some time. The outcome will impact not only upon computing but also upon some of the most vital academic and administrative aspects of the institution.

A few examples highlight the range of alternatives available for providing computing in higher education: internal and external networks; the emergence and effect of the mini's; the Chi Corporation; and TUCC; various cooperative ventures, and the self-sufficient institution. These alternatives will be described in the Denver Study final report. However, the alternative does not alone determine the environment in which computer decisions are made. The following games are generalizations of games that have been observed in one form or another at several institutions.

At the October, 1972 meeting of EDUCOM at Ann Arbor, Professor Berg discussed networks in terms of economic trade theory.[5] Our first game is based heavily upon Professor Berg's paper and is called the "Balance of Payments Game." This game is played by several institutions who decide to share their computer resources through a network. The basic idea is that each institution will offer unique services to the network. The individual user at an institution has a free choice of systems. The institution guarantees to pick up the tab for machine time used on other participating systems and associated network costs. This seems like an ideal situation. The idea behind this concept is that each institution can develop its own specialized facilities and software and let other institutions develop other capabilities according to its own special talents and a free exchange of trade will occur. However, the goal of the game that is being played is to maximize income. Institutions really want to derive more income from the selling of services to other institutions than they will spend buying services from other institutions for their own faculty, students and staff. In fact, every institution in the game is probably trying to maximize a positive balance and minimize the possibility of a negative balance. Since there is a wide diversity in the pattern of outside support of disciplines and the types and machine requirements of services used, a balance of payments problem may develop. Soft money previously used to pay for computing done at an institution may well become hard money leaving the institution. This may result in imposition of tariffs and the subversion of the educational process to produce marketable computing services rather than services of educational value. A game is often also being played even if one avoids the problems of tariffs and subversion. Often the institution must fundamentally change the way it allocates and budgets money. For example, many institutions do not allow a department to carry a deficit or surplus forward from year to year. In fact it is pretty hard for an institution to carry such balances into future years. Only if institutions change to a long term budgeting process for computing can they hope to achieve a balance of payments.

The "High Society Game" also involves cooperative ventures but is usually played by administrators in, the so-called, prestigious schools. These administrators feel that they would rather do their computing with schools of "similar academic and social standing" rather than join together with less prestigious institutions even though the latter may have the capability of providing better or lower cost service, but with lower prestige value. Make no bones about it, the computer is a symbol of prestige. The institutions that play the high society game, more often than not, are private and the mere thought of a student running a program on a public institution's computer conjures up visions of the great unwashed. There is no way for some large, diverse public institutions to understand why certain small "high prestige" colleges turn down their offers of help unless they understand the game isn't service, but high society.

The inverse of this game is the "Welfare or Proud Poor Game." This game is played by a large public institution truly trying to serve the needs of the state and making its services available to the other, poorer state institutions usually including four and two year colleges. Of course the time made available is excess time and the hope is that the meager income derived will help pay for an overly large system. It is also hoped that the services so generously rendered will be looked upon with favor at the next legislative budget hearing. This will enable the providing institution to buy even more excess capacity and staff. The recipients of the largess react as the proud poor saying, "If you don't want to give me adequate services, equal in quality to yours; then I don't want it. I will find my own services and pay for them with my own resources, even if they are of a much lower quality and quantity. At least the computing resources will be mine." It will also increase the receiving school's prestige value. Only when one understands the "proud poor" can one understand the low usage of some welfare networks.

The "Faculty Power Game" is an interesting one and is played at large institutions with a strong tradition of faculty governance. At one institution where this game was being played, the administration was faced with a proliferation of mini's on campus (where a 360/50 may be considered a mini). The institution also had faculty members capable of obtaining outside support with hard money to buy computing equipment for their own research projects. The administration decided to appoint a faculty committee to oversee all computer operations and acquisitions.

The intent was noble: get the faculty to channel the money into the main systems which were serving a declining clientele because of the competition. However, the administration didn't realize that the game being played was faculty power first and computing power second. The committee ended up advising, and even re-writing, a faculty member's request for a system of his own so that it could be justified to the institution. This game in the end, attempts to maximize faculty power and minimize administrative power, regardless of the merit of the computing decision.

The "Godfather Game" is played by the entire institution. The Godfather is the key person on campus who really controls computing and to whom all computing decisions are referred whether or not he shows up on the organizational chart. The Godfather has been a faculty member since 1940. In the mid 1950's he was responsible for bringing the first computer on campus. He programmed it, ran it and taught others to use it. He was responsible for the institution's first (and also the second and third) N.S.F. grant to buy a larger system. He may have moved on from computing center manager to start the new Computer Science Department. He is now, in his words, "only a user of computers." But make no mistake about it, he is the Godfather of computing at the institution. No decision is made without consulting him. If something is proposed to the administration, the question is, "What does the Godfather think?" If one wants something done, one doesn't see the computing center director or the computer advisory committee, one sees the Godfather.

There are many other games that are played – the "Brokerage Game" and the "Financial Boss Game" for those with a knack for handling money, the "Colonial Expansion Game" and the "Fiefdom Game" for those with an imperialistic bent and the "States Rights Game" which can be summarized by the threat, "If you don't go along, you'll get your computing from the state game and fish department."

The issue of selection and evaluation of computing alternatives boils down to how the institution really makes decisions about all matters of a pressing nature. Decisions regarding computing and computing alternatives involve an issue that many top administrators are unable to deal with. They are forced to make decisions about something which, by-and-large, they know little about. Because the real information is in the hands of those who have vested interests, the game starts. Administrators are not inept or ineffectual. By-and-large they understand the alternatives and the implications of choosing a particular alternative. However, they look to their specialists for technological information. Computer specialists must often try to present a complex but technologically and financially sound idea to the administrators in an institution. Many administrators understand, even if they don't admit it, the viability of the plans. However, the heart of the matter is the particular character of institutional dynamics at work. It is these dynamics that really determine how the decision is made. Most computer alternatives are proposed to the administration couched in technological terms and jargon. However, computer decisions are made the same way other major decisions are made and are subject to the same constraints. Because selection of a computing alternative will often change the institution, one must first understand the institutional dynamics involved and then work within these parameters in order to play the game and win.

REFERENCES

1. N.S.F. Grant GJ32368.

2. Greenberger, Martin, "Computing in Transition", *Science*, Vol. 181, No. 4106, 28 September 1973.

3. There might possibly be a larger but definitely separate report containing the supplemental and more technical information.

4. Berne, Eric, The Games People Play (New York: Grove Press) 1964.

5. Berg, Sanford, "Networks in Economics", *Networks and Disciplines*, (EDUCOM Princeton, N.J.) 1973.

The Information Utility in a University Environment

by Paul Stephan*
Case Western Reserve University

THE NATURE OF INFORMATION PROBLEMS

Why should one attempt to consider information, or to be more precise, information systems, as a utility type of enterprise? Two reasons are apparent. First, the nature of information problems existing at present and anticipated in the future requires solutions that are considerably more user oriented than current approaches. Second, the precedent of successful operations of public utilities as user oriented enterprises suggests that the concepts developed in existing with utilities may be applied to operations of some information systems. Let us first consider aspects of information problems that invite a general reorientation of our thinking about possible solutions.

The nature of modern information problems can be perceived as a quantity/quality paradox: overabundance of information at the source ("information explosion") and scarcity of relevant information for the user at the destination. In an attempt to resolve these problems, many types of information systems evolved each with its own specific use of information technology. However, judging from many user reactions, one must conclude that the user's problem of getting relevant information has not been significantly alleviated. Even worse, it seems that many solutions based on the new information technology have resulted in new barriers, imposing new levels of complexity and difficulty for the user. There have been successes as well. Certain technological advances have realized the promise of reduced clerical efforts, decreased likelihood of clerical errors, and rapid physical transport or retrieval of information. Concurrent with these advances however, is the realization that information problems are by and large not technologically limited. Solutions to information problems have simply not yet defined what it is that technology should do.

---

*Dr. Stephan presented this paper at the EDUCOM Fall Conference in place of Dr. William Goffman, Dean of Library and Information Science, Case Western Reserve University.

As one looks to the future, certain problems stand out for attention. First, the increasing complexity of social structures and problems places an increasing need for information. However, the ability of information systems to satisfy complex requirements is not increasing at a rate sufficient to keep up. Second, in highly developed societies everyday life is becoming more complicated and individuals, increasingly need all kinds of information to function. Furthermore, as one becomes more and more assured of equality in goods and services, one starts to seek new services related to quality of life and new modes of participation in social processes. For all this a new type of information system is needed that is specifically oriented toward the general populace, not only toward the intellectual elite. Third, the human social order is in a transition period from the industrial to a post-industrial society; there is an increase, almost to the point of dominance, in types of work requiring higher education. Knowledge is becoming the most important social force and knowers, professionals, technicians, and managers, the most important social group. As a result, communication and management of knowledge will be one of the most essential areas of the new post industrial society. New concepts will be required for many types of information systems. Finally, policy making and management related to environment may become a most important factor in the survival of our civilization. Since rational policy making and management is based on information, one may expect an increase in demands for special kinds of new information systems related to policy making and management.

Obviously all of these problems and demands will not be resolved simply by approaching information systems as utilities. However, utilities may provide a conceptual framework and general model around which some information systems may be organized, and some critical information problems alleviated. A number of proposals to this effect have been made.

In summary, the recommendation for considering information systems as information utilities rests on two primary expectations. First, on the level of psychological impact, the concept of a utility type of enterprise evokes a positive response because it is associated with the elusive but important idea of operating for public good. It is hoped that this will become the major orientation of information utilities as well. Second, on the level of performance, the particular enterprises which are recognized as public utilities indeed are quite successful in dealing with problems of regulated supply of commodities or services with unique characteristics. It is hoped that information utilities will achieve a similar success in performance.

INFUT

This paper reports the progress of an investigation named INFUT* (INFormation UTility) at Case Western Reserve University, which is attempting to determine the conditions under which a wide variety of information sources and systems could be considered as a utility. The Project INFUT addresses information problems at the organizational system structure level. That is, the "software" needs revision, not the "hardware". The approach is to structure the information source/user relationship as a utility; not unlike public utilities that appear for power, natural gas, and the like. Such a structure offers a powerful new philosophy to approach the complex informational problems we face today. The specifics of the investigation are limited to information systems within a university. However, generalizations may be made to other environments. Specific areas investigated include: 1) the nature of present and expected information problems that invite consideration of information utilities; 2) the characteristics and properties of public utilities from which the characteristics and properties of information utilities can be inferred; 3) the factors in the nature of information and properties of information systems that will affect development of information utilities; and 4) the attempts to specify those properties of information utilities that appear within the environment of a university. This paper concentrates only on two of the aspects investigated in Project INFUT: 1) properties of public utilities; and 2) implications for information systems. Other aspects of the project are mentioned but are not elaborated.

The objectives of the research are to provide general answers to the following basic questions:

- What are the properties of public utilities that may be useful for the design and operation of information systems as a utility?
- How can the use of all information systems and sources in a university be viewed in a unified way that will allow for isolation of the elements and study of the interactions essential to the operation of a utility?

---

*Work performed under the National Science Foundation Grant No. GN-36085, Tefko Saracevic, Paul Stephan, and Douglas H. Rothenberg principal investigators.

- What properties of information sources, users, and systems in a university lend themselves to the utility concept, and what properties do not?
- How does one design an information utility from the ground up?

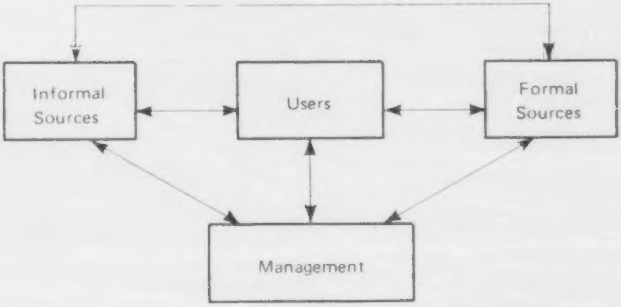
Two approaches were taken in relation to each of the preceding stated objectives. First, a study of public utilities, including a review of the literature and interviews with several levels of management, led to a conceptualization of the generic properties of existing utilities. These properties are discussed later in some detail. Using these generic properties as a guide, a general utility notion was constructed. It is this generalized utility notion that serves as one facet of the "starting point" for the exploration of information as a utility. Second, a study of existing paths of communication (written, media, and oral) within the university was made using current theories and models of communication theory in order to produce a somewhat general model of the utilization of information in a university environment with a primary orientation toward the user and, potential user. Additional work was intended to sharpen the general notions into concrete deductions or inductions. A series of empirical studies were initiated in order to obtain quantitative evidence as to: 1) the validity of the general model mentioned above; and 2) the types of interactions that do or do not tend to characterize existing information components in the university as a utility. The final part of the objective is yet to be done. Using the previous studies (above), practical methodologies for the study, design, and development of information utilities will be developed.

THE GENERAL MODEL

Any university is a conglomerate of information systems, hopefully possessing primary functions of teaching (communication of knowledge) and research (search for new knowledge and organization of knowledge). Elements of this conglomerate may be studied in isolation, but for a unified view (a common classification scheme, a general model) it is necessary to study interactions in a consistent manner. Moreover, a unified approach may uncover sources and effects quite different from those commonly assumed. The unified picture used for this work evolved from considering the Shannon-Weaver model of communication process, combined with Wiener's cybernetics and Goffman's epidemic theory. However, these were first modified by placing the users as the central element in the conglomerate of systems because users are the focus in the establishment of a utility.

Four basic classes of interacting elements have been defined: users, formal information sources, informal information sources, and management. Users and potential users include students, faculty, staff and administration, and outside users. Formal information sources include curricula, library collections, AV/TV and media collections, and computer data. Informal information sources are invisible colleges, and personal collections. Management, including policy making and allocation of resources, occurs at the level of each individual source, at the level of users, and at the university-wide level. Schematically, these four classes and their interactions are represented in figure 6.1.

Figure 6.1 Schematic Structure of the System



Empirical studies conducted during work on project INFUT have assumed that the key for consideration of an information utility is the interaction between and within the various elements (classes and subclasses) enumerated above. Thus, the empirical studies have been oriented toward illuminating the various stages of interactions. Inventories of factors in possible interactions within each element have been made, and interaction between elements from two, three, or four different classes or subclasses have been initiated. The following studies are in progress, and others are planned:

- Library use study: circulation-demands over time; demands by users; characteristics of non-users; un-met demands; demands by subjects in relation to curriculum.
- Informal sources: private holdings of faculty as one type of informal source, extent, nature, investment, and availability; relation to formal sources.
- Computer use study: facilities; nature of use and users; demands and un-met demands; relation to other formal sources.
- AV/TV use study: facilities; nature of use and users; demands and un-met demands; relation to other formal sources.

- Faculty research study: pattern of relations between faculty on the basis of their publications, as opposed to departmental organization.
- Faculty teaching study: pattern of relations between the various faculty on the basis of the sharing of students.
- Curriculum study: pattern of relations between subjects in the curriculum; demands on curriculum; access to curriculum.
- Management study: nature of regulations imposed; pattern of policy making; information on the basis of which policy and allocations are made.
- Architecture and geometry of the general utility.
- Growth and decay of utilities.
- Historical perspective of utilities.

The Future of University Computing Centers

by William Kehl
UCLA

Some old saws persist today about the services offered by university computer centers. Many campus computing centers were started by faculty who were not skilled in management, who became entrepreneurs and provided poor services coupled with financial irresponsibility. Unfortunately, this worn-out image persists and, from this premise, the crystal ball deduction is quickly made that the days of university computing centers are numbered.

The trouble with this image is that it has not been valid for some time. University administrators ought not to put reliance on this old story as a way out of present fiscal problems. It was true that faculty who were seduced into starting computing centers on their campuses around 1956 became enamored with their new status, but that era is long past. The professor who doubled as a part-time computing center director has often found a home in the new Computer Science Department. In many ways that it what he really wanted all the time.

University computing centers have matured. Center directors today are very conscious of services and of fiscal responsibility. Often campus centers provide a variety of good services and unique capabilities to respond to academic needs that no commercial service can match. Center directors today carefully manage computing resources and are supported by a close organization of policy committees with fiscal controls which oversee services. Oddly enough, the federal funding agencies have not recognized this evolution. Grant review committees are still staffed exclusively with the same entrepreneurial faculty. The newer, responsible computing center manager who is in tune with satisfying academic needs does not yet play a significant role in grant review committees or in national plans for support of computing needs in research.

Current problems in computing center operations are due not to mismanagement, but to outside factors. By the late 1960s, most campus center operations had begun to take on an orderly pattern. Since then, centers have been subject to the same cutbacks that other university programs have suffered. In most cases, overexpansion beyond campus needs has not occurred. With the general cutbacks in university funds, campus centers just are not able to meet the needs. In addition, they have been plagued by inconsistencies in government practices and funding.

FEDERAL FUNDING POLICIES

Universities are subject to strict government audit of expenses for provision of computing services on government contracts and grants. Campus computing centers have to be cost effective and have to recover *all* costs in an equal and fair manner for both research and instructional use. To conform with current audit regulations, records must accurately reflect this balance. On the other hand, the government does not follow a full cost recovery policy. Computing at the AEC National Laboratories, for example, is artificially subsidized because equipment costs are capital purchases that are not included in rates of charge. Thus, rates are half of what they would be under a full cost-recovery system. Because the National Laboratories do not have to recover full costs through rates charged for services, they can undercut any university computing center, no matter how efficiently it is managed. Seventy percent of Lawrence Berkeley Laboratory's computing in recent years has been from outside work. However, the AEC is not the only government agency which writes off the capital cost of the equipment. Other federal agencies have been lax in calling to account computer services derived from direct equipment acquisition under grants and contracts. Under these policies, National Laboratories and agencies have little incentive to be cost effective. Under the NSF Institutional Computing Services program between 1956 and 1969, grants for computer purchase helped many universities meet the responsibility to provide computing services by reducing the risks to the university. A grant represented an investment by the government in the university's contribution to research. The ICS program contributed more than any other, before or since, to the advancement in the quality and quantity of university computing service available for research. With surprisingly few dollars it was able to meet campus computing needs because university programs for research and instructional academic computing were better coordinated and designed to be mutually reinforcing. The ICS program was unfortunately discontinued in 1970.

The government must now take more imaginative steps to overcome inconsistencies in funding policies and to help relieve the constant thrashing that is going on due to the unstable funding levels. One step might be to make *all* centers both in universities and government laboratories total cost-recovery operations. If all centers had to cover capital acquisition costs, equipment purchased with federal funds would be amortized and users would pay full charges for services. In this environment, computer networking could develop in a logical pattern. Today it is often considered a panacea for those who wish to escape local financial problems and commitments.

Another area where an imaginative approach on the part of the federal government is necessary is in understanding the relationship between responsibility and financial risk in a university. A university takes considerable responsibility and, hence, risk for its faculty to provide the general academic environment, including laboratories, libraries, other research facilities, and computing capabilities. If these services are to be provided through a coordinated plan, computing centers must evaluate carefully responsibilities for instruction and research. As they try to provide facilities for the faculty as a whole, universities must consider the increase in risk due to federal grant policies. It seems as if universities must take all the risk and the government takes none. Yet, if a campus cuts back on computing resources, and leaves to chance the possibility that the funded researcher can find on his own the services he needs, risk is also incurred. A responsibility is avoided.

New solutions are needed. However, of basic importance is the necessity for the federal government and the universities to work more closely together on these problems.

Chapter 7

Considerations of Networking

Network Management for Expanded Resource Sharing*

by Dennis W. Fife
Institute for Computer Sciences and Technology
National Bureau of Standards

Remote access computing via data communications facilities has grown tremendously over the past decade, and is now established as an important future growth area. The recent connotation of resource sharing has been fostered by Government-funded projects emphasizing either computing service distribution and user support, e.g., NSF regional networking project[1], or innovative communications facilities, e.g., ARPANET. The future availability of communications capabilities for resource-sharing is very promising. A variety of techniques and systems have emerged operationally successful[2], and commercial development activity is high, e.g., the newly-announced ventures to commercially exploit the ARPANET technology. On the other hand, the future is uncertain regarding cooperation among autonomous institutions and expanded support for research computing capabilities that are not commercially appealing. Significant cost savings may not develop without large-scale service aggregation and reorganization that normally are achievable only under strong central management. Also, remote access tends to aggravate the user's operational problems unless there is additional aid in documentation, etc. Thus important issues and research problems exist in broad areas of network management, remote user support, and associated software management practices. The new NSF research initiative in networking[3] addresses these unresolved problems through research studies and experimental network operations.

---

*Supported by the National Science Foundation, Grant AG-350. Contribution of the National Bureau of Standards, not subject to copyright.

The early regional networks were dominated by one major institution which naturally provided central management. A highly significant distinction in nationally-oriented experiments will be the collaboration of several major institutions, each with appreciable computing capability to offer. Moreover, financial investments may be advisable to supplement research grants in initiating an experimental network. Future networking experiments must then involve high-level management initiative and careful negotiation to form an acceptable management organization.

A carefully prepared network management plan is one of the most important elements of a networking experiment, helping to estimate the potential success of the experiment and to guide other experimenters. To encourage innovation and to accommodate the varying needs of participants, it is not advisable to delineate one management structure as uniformly desirable. Even so, the organization and characterization of primary management issues as they might be addressed in a management plan is the purpose of this report. This results in a preliminary evaluation mechanism for network management, insamuch as it aids an initial review of completeness and pertinence, prior to a detailed evaluation of the specific considerations of an experiment.

NETWORK MANAGEMENT FUNCTIONS

Network management encompasses four general directive functions. These may be identified by reviewing the management organizations of existing network operations[4], and also by extending the traditional features of complex features of complex service organizations.

Policy Direction. Participating institutions will require continuing review and directive authority to ensure that goals and policies in a networking project are consistent with their objectives, philosophies, and requirements. A common approach is to establish a policy committee or a board of trustees composed of high-level officials (agency director, dean, or provost) who represent institutional responsibility for the project. The policymakers preserve control over the means to satisfy their institutions's computing requirements.

Executive/Technical Direction. Translation of general goals and policy into networking accomplishments requires leadership to formulate specific objectives and approaches, to distribute resources, and to monitor progress against required performance. Ideally, executive direction would be assigned to one individual, but co-equal directors and committees have been accepted compromises. Having overall responsibility for the technical facilities and objectives, the executive director is motivated to obtain sufficient resources and support, and to meet participant requirements expeditiously within the available support.

Operations Direction. The technical complexity and workload usually involved in the daily provision of remote access service demands a skilled and responsible individual to manage the operations and maintenance staff. Although each computer center will have an operations manager, coordination problems and the interaction with data communications operations may require a "network" operations manager. The principal objective would be to maintain high reliability and service effectiveness within assigned resource capacity. Also, operations should help identify unfulfilled requirements for future development or resolution.

User Assistance/Marketing Direction. From both an administrative and technical standpoint, remote users require a basic amount of assistance in gaining access to a service and some continuing aid in solving operational problems. As a very limited and easily consumed resource, the available time and skills of consultants, aids, and training specialists must be judiciously assigned to scheduled and unscheduled activities. The objective of the "marketing" group is to develop user awareness and dexterity regarding basic operating procedures and available services, within their capacity.

PROGRESSIVE STAGES OF RESOURCE SHARING

When the scope of resource sharing being undertaken is described, the above definitions can be extended to specific concerns for each function. However, past identified types of resource sharing – e.g., load sharing, data sharing, program sharing – are not particularly helpful since they relate more to software than to management. Five new categories are proposed here, describing evolutionary stages that may help to understand and evaluate network management requirements. Each is a necessary stage preceding its successors, although it is apparent that a network could be implemented to pass over one or more stages concurrently.

Mutual service access. This initial phase of resource sharing merely establishes possible access by individual users to multiple remote computer systems. Institutions thus permit all cooperating external services to compete with their local computer center to meet their user's needs – but perhaps not in free competition. Policy modifications and management action may be necessary to revise outside procurement rules and financial procedures. The user would deal directly with each available service regarding his needs. There would be minimal management interaction with users or computer centers.

Mutual support. This phase develops when each organization establishes formal assistance regarding accounts, documents, training, and usage problems, to aid its members in using all available computer systems. Some mutual support in financial transactions, e.g., billing and collection, would be appropriate also. The independence of computer centers would be preserved.

Operations coordination. This phase requires organizations to arrange operational or management procedures and criteria in agreement with other centers, to provide uniformity and ease of user access to various systems. Common accounting criteria, protocol elements, and control/communications factors (e.g., prompting characters) would be typical considerations. The resources offered by each computer center would be of its choice, but demand would change with easier access and use. Explicit methods for load sharing and balancing might be introduced at this stage, but each computer center would be expected to have independent discretion in setting priorities with respective to the market.

Service alignment. This stage arises when organizations mutually recognize the resources each can best provide, and legally apportion their overall market among themselves. As a result, computer centers would specialize in certain types of service, and organizations would become mutually dependent in satisfying their composite needs. Such dependence would be significant financially, for an institution may require an assured market to support its investment in offered services.

Joint resource ownership. This stage evolves when organizations mutually invest capital and personnel resources to develop a new computing resource for their common use. An institution would thereby acquire a vested interest in another organization's computer center. As a result, participants become most strongly committed to resource sharing through joint ownership, and the operating independence of joint computing centers is more constrained than in other forms of resource sharing.

NETWORK MANAGEMENT RESPONSIBILITIES

The preceding definitions give a two-dimensional perspective on the tasks and goals that should be addressed in a network management plan. The following table briefly characterizes the principal developments or products expected from management activity in each case. In evaluating the proposed management for a networking experiment using this guideline, the basic questions become: What organizational entities are responsible for each development? What specific form does each product take? Are all objectives considered and what secondary tasks are identified? Other NBS analyses[5] would help in the detailed consideration of these questions for technical tasks in Operations Direction and Executive/Technical Direction.

University computer networks today seem largely to represent resource sharing at the stages called Mutual Service Access and Mutual Support, although the Triangle Universities Computing Network[6] perhaps serves as an example of the Joint Ownership stage. In the advanced stages, there are substantial technical and management problems to be solved, especially ones affecting economic viability of networking and service quality for a broad user market. The needed techniques for market research, quality control, and economical service packaging, and the feasible scope of standardization for software and operational procedures draw particular attention to the Executive/Technical Direction function. In the present evolution of resource sharing, networks may operate successfully without distinct recognition of this role – technical direction may be distributed among various university officials and the end user. But as Operations Coordination emerges, forceful leadership seems essential in recommending and implementing uniform operations, e.g., standard log-on sequences or pricing algorithms. As cooperative networking progresses to the Service Alignment or Joint Ownership phases, the need for a strong Network Manager seems clear. As in a commercial enterprise, this executive must evaluate the trade-offs perceived by different contributing organizations, and present the policy group (probably a board of distinguished trustees) with a balanced choice in quality services and associated development projects.

Figure 7.1 categorizing network management responsibilities, and a forthcoming final report[7] on the NBS network management study, are meant to draw attention to the critical importance of a carefully conceived management pian. NBS is distributing this and other reports to participants in the EDUCOM Working Seminars and recent conferences. Other interested individuais may write Dennis Fife at NBS, to receive the recent reports and to be included in distribution of future reports which will emphasize service quality assurance and measurement.

Figure 7.1 Principal Developments of Network Management

MUTUAL SERVICE ACCESS	Acceptable needs for external service, and the expected impact on the local computer economy.	Consolidated reporting of computing needs, and adaptation of funding practice to procure external service.	Scheduling and essential operational improvements for remote service.	Economical assistance and marketing to remote user communities.
MUTUAL SUPPORT	Acceptable scope and funding for joint support arrangements, to meet user needs but stabilize external service demand.	Budgeting and technical guidance for balanced, effective support activities.	Competitive system improvements and enhancement of remote support capabilities.	Effective joint support activities with innovation in automated user aids.
OPERATIONS COORDINATION	Acceptability and priority of operations standards, to meet user needs with economic feasibility.	Recommendation and implementation of uniform operations plans and standards.	Competitive proposals of simplified operations; implementation of uniform interfaces and procedures.	User needs for and satisfaction with uniform operations and service access.
SERVICE ALIGNMENT	Selection of primary and alternate resource suppliers, for quality services with competitive innovation.	Formulation of service allocation concepts; marketing research, and evaluation of supplier capabilities.	Formulation of service offerings; evaluation of resource viability.	Aid in integrated applications of distributed, complementary resources.
JOINT OWNERSHIP	Criteria for investment participation; selection of resource development projects.	Evaluation and recommendation of candidate projects; technical and fiscal management of overall development.	Formulation of proposed resource responsibilities; development and operation under assigned conditions.	Needs analysis to define new resource prospects.

REFERENCES

1. F. W. Weingarten, et.al., "A Study of Regional Computer Networks", University of Iowa, September 1972.

2. R. P. Blanc and T. N. Pyke, Jr., "Computer Networking Technology – A State of the Art Review", *Computer*, August 1973.

3. See D. D. Aufenkamp, "NSF Activities Related to a National Science Computer Network", *Computer Communications: Impacts and Implications*, Proceedings of First International Conference on Computer Communications, Washington, D.C., October 1972 and, M. Greenberger, et.al. (Editors), *Networks for Research and Education: Sharing of Computer and Information Resources Nationwide*, MIT Press, Cambridge, 1973.

4. See I. W. Cotton, "Network Management Survey", NBS Technical Note 805, February 1974, and A. J. Neumann, "Review of Network Management Problems and Issues", NBS Technical Note 795, October 1973.

5. Several references are relevant: A. J. Neumann, "User Procedures Standardization for Network Access", NBS Technical Note 799, October 1973; R. Stillman, "Computer Networking: Approaches to Quality Service Assurance", NBS Technical Note 800, January 1974; R. Blanc, "Review of Computer Networking Technology" NBS Technical Note 804, January 1974; A. J. Neumann, "Network User Information Support", NBS Technical Note 802, December 1973; and R. Blanc, "Cost Analyses for Computer Communications", NBS Technical Note, Expected publication May 1974.

6. See L. H. Williams, "A Functioning Computer Network for Higher Education in North Carolina", *Proc. FJCC*, 1972.

7. D. Fife, "Research Considerations in Computer Networking to Expand Resource Sharing", NBS Technical Note 801, April 1974.

An Information Dissemination Network Model

by Glenn O. Ware
and
John H. Schuenemeyer
University of Georgia

Over the last five to ten years a number of computer-based information centers have been developed which, along with the existing library structure, have greatly increased user accessibility to available information resources. These centers, many of which are university-based, have adapted the use of computer resources and to some extent telecommunications technology to the problems of providing access to the large bibliographic and data resources which have traditionally been handled by libraries. The developmental activities of these centers have been largely autonomous with each center attempting to solve the technical problems associated with its own environment.

With the increasing pressure to continue to provide and expand the delivery of information resources to the scientific community, a national science network composed of the existing libraries and information retrieval centers seems to offer the most logical and natural means of doing so on an economical scale. Working relationships between various centers have been established during the past couple of years. However, the arrangements have been established on an experimental basis with no permanent economic mechanism being created for resource sharing among the centers.

As we look to the future and focus our attention on a national science network, our goals should be directed toward providing the transfer of information resources between science and society under viable economic conditions. During the past year, the University of Georgia Information Dissemination Center and the Center for Information Science of Lehigh University have jointly investigated, with support from the National Science Foundation, alternative approaches for the possible integration of computer-based information dissemination centers into a network.[1] As a result of this investigation, a simulation model has been developed representing the operational characteristics of an integrated information network. Our principal objective with the network model is to provide statistical information on network activities to allow one to explore and evaluate various alternative approaches to network implementation. Although the model is oriented toward computer-based information retrieval centers providing search services for bibliographic data bases, the basic concepts involved in the development of the model should have a wider range of application.

COMPONENTS OF THE INFORMATION NETWORK MODEL

The information dissemination network model represents the integration of a set of information processing centers whereby users can interact with the centers through a common communication system to retrieve desired information from any number of data base files. The model divides the various functions of the network into four basic components which include: 1) users; 2) information resources, such as the bibliographic data bases; 3) information processing centers; and 4) the communication linkage of 1, 2, and 3 into the information transfer network.

USER POPULATION

Each information processing center has a variety of user groups, and the services provided by each center are structured around the needs of this user population. The user community may encompass not only the faculty, staff, and graduate students at the university-based centers, but also users from commercial organizations, governmental agencies, and other institutions and universities lacking information retrieval resources. Both the size of the potential user population and the geographic distribution of this population are important factors when considering a national science information network.

INFORMATION RESOURCES

The information resources consist of the computer-readable bibliographic data bases which provide the basis for all system transactions. The majority of the files are produced as secondary products by the abstracting and indexing services, though not all files have corresponding printed versions. The files received from the suppliers are normally converted to formats acceptable for processing by the search systems of the various information retrieval centers. The number of computer-readable files has exhibited a relatively steady increase during the past few years. A recent survey revealed that 48 publicly available data bases are being processed either for current awareness (SDI) or retrospective searches[2]. However, the majority (44) of the data bases were used by fewer than ten centers; only four data bases were used by ten or more centers.

The subject areas of the bibliographic files cover a wide range of disciplines and the contents of the files vary greatly, both in terms of the types of documents covered and in the types of information occurring in the files. The types of documents include periodicals and other forms of serial publications, patents, government reports, books, and maps. The information in the files includes various data elements such as titles of publications, authors' names, index terms and classification codes, bibliographic citations for referenced documents, and often abstracts.

INFORMATION PROCESSING CENTERS

As of March, 1972 there were 25 full members of the Association of Scientific Information Dissemination Centers offering bibliographic search services[3]. Most of the centers have either developed their own generalized computer software system to support the search services or have adapted software packages available from some other source, such as computer hardware vendors or commercial software houses. The search systems of these centers can be divided into two major types of processing. In batch processing all profiles coded for a particular data base are selected for searching in batches. On-line, interactive searching involves a direct computer-user interaction with the user making queries against the data base file. Several users may be interacting simultaneously with the data base file.

Regardless of the processing philosophy, the operational functions of the two types of processing centers are very similar. Both centers acquire computer-readable information files, convert them as necessary to formats suitable for processing in their own environment, update their retrospective collections if the retrospective searching of files is provided, search the user queries against the file, and disseminate information from the files to individual users. Of primary importance in a national science network is not only the functions which are to be performed by the member processing centers but also the resource requirements of these functional activities for meeting the anticipated demand.

COMMUNICATION SYSTEM

The possibilities for the interconnection of the network nodes range from mail service to sophisticated computer-based communication facilities. Several modes of communication can be employed in any given network, with the choice of mode being dependent on network size, volume of information being processed and delivered to different nodes, desired turnaround and response time, costs, and technical feasibility.

The communication system for the interconnection of the nodes is evaluated in terms of the information flow requirements like estimates of the amount of information needed to be transferred to where and at what time. A major goal in these evaluations is establishing how information flow should be managed to achieve operational efficiency, stability, and economy.

The information flow in the information dissemination network is usually considered bidirectional, with search requests and associated messages directed by the user to the processing center and the search results (bibliographic citations) directed from the processing center to the user. The volume of information transmission in terms of the number of characters transmitted from the center is significantly greater than the number of characters transmitted by the user.

REPRESENTATION OF INFORMATION NETWORK ACTIVITIES

In order to provide an accurate determination and representation of the network nodes and their interactions, data was gathered on the operational activities of two typical information dissemination centers: the University of Georgia Information Dissemination Center and the LEADERMART Information Service Center of Lehigh University. The search systems of these centers are characteristic of the two major types of information processing philosophies. The University of Georgia bibliographic search system features an on-line data entry and maintenance system for search requests and user information with the retrieval portion of the system operating in batch mode[4]. The LEADERMART search system is an on-line, interactive technical information retrieval system[5].

The University of Georgia Center has been in operation since 1968 and data on the operational activities associated with the search services for 17 separate data bases have been collected since the initiation of the service on each data base. Data on the operational activities of the LEADERMART system were collected for a 107 day period. From statistical analysis of the data, quantitative measures and prediction equations have been derived and incorporated in the network model for estimating various parameters associated with user access demands, data base holdings, computer processing requirements, and communication needs. Listed below are examples of some of the network operational activities which can be accurately predicted:

- The number of messages and corresponding characters associated with user access requests.
- The central processing unit (CPU) time requirements for the conversion of data base files for computer processing.
- The CPU time and core memory requirements for processing varying levels of user requests.
- The number of characters associated as output for the user requests.
- The traffic flow requirements between network nodes.

SIMULATION OF THE NETWORK MODEL

The actual simulation of the model projects the operational activities of the network through a planning period of some given time period. A considerable number of input parameters are used by the model in performing the required projection. For a complete description of the input requirements of the model, the interested reader is referred to the report by Ware and Schuenemeyer[6].

The simulation model allows one to examine the effects of varying the input parameters on the computer processing requirements of the information dissemination centers and the telecommunication requirements to handle the resultant transmission loads. The effects of changing anticipated usage levels and information resources growth levels can be examined. Also, the effect of varying the distribution of the data bases among the centers can be considered.

Several types of information processing centers can be examined ranging from a single processing center providing a single type of information service to multiple processing centers, each providing several types of information services. Various network topologies for the interconnection of the nodes are evaluated with the information flow in the network being routed to achieve operational efficiency, stability, and cost-effectiveness.

CONCLUSIONS

The simulation model provides a means of exploring and evaluating alternative approaches for implementing an information dissemination center network. Network analysis performed using the model is strongly dependent upon the assumptions used when specifying the input parameters and input data for the various information processing centers, bibliographic data bases, user communities, and the communication system for data transfer. However, the information dissemination network model is as realistic as possible since it is based upon the actual operational activities of two dissemination centers. The design of the network model is generalized sufficiently to represent the operational activities of similar types of information dissemination centers. It is hoped that the model can be useful in planning for the development of a national science information network.

Several areas which need additional investigation include: 1) procedures for scheduling network activities in terms of traffic flow and processing load requirements of the information network while maintaining adequate user response time; 2) procedures for optimizing the allocation of data bases and/or partitioned subfiles among the information processing centers based on user access demands; and 3) a user interface for a multidisciplinary bibliographic information network to enable users to interact readily with heterogeneous retrieval systems.

It is of special interest to see if, by using scheduling and optimizing procedures, large numbers of users may be serviced by an information dissemination network using low speed communication lines like voice grade lines. Such a specialized network may be appreciably cheaper to operate for the participating information centers than functioning within a general purpose network. Preliminary investigations point toward this conclusion, but additional investigation is needed.

As previously mentioned, the development activities of the bibliographic centers have been largely autonomous. As a result, each center's search system and its approach to servicing its user community tend to be different. The extent to which these differ, especially in terms of translation among the various search systems, is an important consideration for networking. Research efforts are beginning to be directed toward the development of user interface procedures which will allow translation to all individual systems and, thus, provide a realistic capability for resource sharing among information dissemination centers.

REFERENCES

1. National Science Foundation Grant GN-32214

2. Williams, Martha E. and Alan K. Stewart, *ASIDIC Survey of Information Center Services: Report of the Association of Scientific Information Dissemination Centers*, ASIDIC (1972), 117 p.

3. Ibid.

4. Park, Margaret K., *University of Georgia Text Search System — System Description and Program Documentation*, (Computer Center, University of Georgia), 1972.

5. Hillman, Donald J. and Andrew J. Kasarda, "The LEADER Retrieval System", *AFIPS* Vol. 34, (AFIPS Press, Montvale, N.J.) (1969), p. 447-455.

6. Ware, Glenn O. and John H. Schuenemeyer, *A Simulation Study of an Information Dissemination Center Network, Part I,* (Office of Computing Activities, University of Georgia, Technical Report UGA/OCA-73-1), 1973, 128 pp.

The Economics of University Computer Networking*

by D. A. Dunn
Stanford University

Many university computing centers have begun to find that the facilities they offer to users are not adequately matched to the demand for services generated by on-campus users. Legal constraints ordinarily limit the sale by universities of surplus time to commercial users and it is risky to invest heavily in new facilities when the market for services is both constrained and made uncertain by variations in the federal funding of research.

Three major alternatives are open to universities in this situation.[1] First, an attempt can be made to match supply and demand locally, following a conservative policy which avoids excessive risk at the cost of only being able to offer second rate computer facilities to local users.

The second alternative is to go out of operation as a source of large-scale computing and require users to go outside for all major needs. Since many commercial sources of computing are now available, this alternative is not unreasonable, but it is often costly because it fails to utilize the subsidies that are available to university computing and the low cost pool of student labor. It also removes a source of student jobs and student training from the university, but it may simply move this source of jobs down the street to the local computer service firm.

---

*Work for this paper was supported by the National Science Foundation under Grant GJ-36392X.

The third alternative is to keep the local computer center alive and to join a computer network which allows local users access to new large computers both at home and elsewhere on the network. Centers with large modern machines gain access to a market large enough to permit them to pay for their machines and which allows them to reduce the risks associated with the purchase of new equipment. This paper is concerned with the economics of this third alternative.

NETWORKING COST SAVINGS

A number of university computer networks are now in operation and many more are planned.[2] The ARPA network is one of the largest and most advanced of these networks and many of its major users are universities.[3] However, networks based on the same basic packet-switching concept are also being designed with commercial users in mind, such as banks, airlines, and hotels. Commercial networks of this type are likely to have somewhat different objectives from those of a university network. In all cases, however, the motivation eventually comes down to cost savings.

Some of the areas in which networking cost savings can be expected are:

- Economies of scale in computation which result from the use of larger computers than would be possible in the absence of networking.
- Economies of scale in data files.
- Economies associated with better use of transmission lines in packet-switched networks as compared with timesharing systems which use conventional dedicated telephone lines between the central computer and individual users.

All of these sources of cost savings are ways of allowing many users to share common facilities and to share the costs.

There are three approaches to the design of networks that allow different degrees of sharing. For example, several user groups that conceivably could share the use of the same network might be airlines, banks, and universities. One approach to network design is to give each user group its own separate network. This approach has the important advantage that the design can be specialized to the needs of the particular user group. For example, the choice of the maximum delay under normal traffic conditions might be quite different for each user group. No compromise is necessary if only one group uses the network.

A second approach is to share only the transmission network, but not data storage facilities or computational facilities. This approach is taken in the ARPA network, and it has the advantage that multiple user groups can share transmission costs without requiring more than a very minimal level of commonality in message format.

The third approach is to share computers and/or data storage facilities, as well as transmission lines. This approach could be the long-term trend, because it offers the greatest potential cost savings. It is an unlikely approach because user groups hesitate, for security and reliability reasons, to share data storage. It is also not clear that cost savings can be achieved by such systems at the present time.

COMPETITION IN THE NETWORK MARKET

Once networking has been accomplished, users will find that they have several alternatives for most types of computing service and computing centers will find themselves in competition with other centers for the business of users throughout the network. As in any other market, price and quality of service will become the determining factors in the user's choice, unless obstacles are put in his way which limit his choice. Computer centers may soon find it in their best interests to form cartels, set prices cooperatively, and divide up the market according to type of service with only certain centers being authorized by the cartel to offer certain classes of service. The pressures on inefficient centers to form cartels or to restrict the freedom of their users will become intense, as soon as services equivalent to their own are available at lower prices on the network. This state of affairs cannot be far away and it is evident that some hard choices will have to be made in the near future.

One aspect of this problem has arisen in recent years in the operation of the Bell System. Although the Bell System has had a legal monopoly in telephone service in much of the U.S. for many years, not unlike the monopolies granted to computer centers by their university administrations, the Bell System has recently found itself facing competition in the long-haul portions of its data communications service. One of the major difficulties it has faced is in connection with its pricing policies. This example may suggest an important class of problems that computer centers may face quite soon operating against competition on a network.

When a firm or computer center is operating as a monopoly, its pricing can be established almost without regard to cost. A monopoly with two products such as residential and business telephone service can, if it chooses, charge more than cost for one service and less than cost for the other, cross-subsidizing one service with excess revenues received from the other. If competitors were to be allowed entry to such a business, it is evident that they would prefer to compete in the service area which is priced well above cost. The area in which Bell is now experiencing competition is exactly this type of service area. For many years Bell has priced long-haul service between two points in proportion to the distance between the points and without regard to the amount of traffic between the two points. Yet, due to economies of scale, Bell's costs were much lower on routes with high traffic densities, so these routes were subsidizing the low density routes. Naturally, its competitors are choosing to compete only on the high density routes.

The pricing rules that a university computer center may have used in its monopoly period may have similar weaknesses which should be corrected when the center goes on the network. For example, pricing computation by the hour, without regard to the amount of core storage used, subsidizes large users of core. Such a policy could lead to overuse by such users, especially if large numbers of such users were suddenly to gain access to a computer through a network.

The simple pricing rule dictated by economic theory is to price each service at marginal cost when faced with competition. With complex services like computation, the definition of a "service" may be rather subtle and needs to be examined carefully. Not only is the overall price of service important, but the detailed structure of pricing of different services must be carefully attuned to the cost of providing each class of service.

LOOKING TO THE FUTURE

The principal economic function of networking is the creation of a market which allows larger numbers of users and providers of computer service to get together. This market may or may not become a highly competitive market, depending on the choices made by participating centers over the next few years. Incentives will exist for the formation of cartels which will both protect the inefficient producers and result in higher prices and a lower output of computing service than if competition were allowed. Pricing will eventually have to be closely related to cost, for all services, if competition exists. If a cartel is formed, however, some services may be subsidized by others.

Eventually the only reason for the continued existence of separate university and commercial computer networks may be legal constraints which limit the ability of university centers to sell computer services in the open market. However, if the university network becomes cartelized and maintains prices well above cost, incentives will be created for users to seek to obtain services from what may then be lower cost commercial sources.

These conclusions are not intended to be predictions of what will happen, but are intended to suggest that university computing centers and administrations will soon be confronted by choices which are not greatly different from the choices that business firms face in meeting competition. Perhaps some of the experience gained in these other areas will be useful in making these choices.

REFERENCES

1. M. Greenberger, "Computing in Transition," *Science*, v. 181, p. 1207, Sept. 28, 1973.

2. D. D. Aufenkamp and E. C. Weiss, "NSF activities related to a national science computer network," *Proceedings of the First International Conference on Computer Communications*, S. Winkler, ed., (IEEE, N.Y.) Oct. 24-26, 1972.

3. L. G. Roberts and B. D. Wessler, "Computer networks development to achieve resource sharing," *Proceedings of the 1970 Spring Joint Computer Conference*, pp. 543-549, (AFIPS Press, Montvale, N.J.) 1971.

Chapter 8

Statewide Plans for Computing

Introduction

by E. Rex Krueger
University of Colorado

Historically, we have seen quite a change in the management of computing centers as represented by the computing center director. This individual generally has assumed his position from the faculty, and more often than not is a major computer user. He has been instrumental in the development of computing within the institution. Often his interest has also resulted in his assumption of the chairmanship of the computer science department. He has been active in soliciting funds, specifically until about 1970, representing the institution to the National Science Foundation by proposal for a facility's grant. The successful director has exploited this opportunity to obtain cost sharing funds from the university and has gone ahead to build a full scale computing resource which often is referred to within the institution as an empire.

In the late 1960's with education generally under critical review, the computing center also found itself operating within an ever increasing set of controls. One of the reasons for this situation was that within the total assignable square footage of the university, the greatest wealth per square foot is concentrated in the computing center. Because the capital equipment represents a significant outlay, the computing center also draws attention. Thus the computing center and its management, having built and economized across the period of the 60's, now finds itself dealing with a new set of controls which undoubtedly have impact on the center's mode of operation.

In some ways this experience is similar to the changes taking place in the automobile technology and the controls on utilization. During the 1960's, cars were considerably improved in terms of the technology involved. They grew in size and horsepower and, of course, gasoline consumption. In the 1970's we began to see new antipollution devices installed. These environmental controls have an impact on fuel consumption which, across a relative short period, contributes to severe fuel shortages. Of course I haven't mentioned the safety control which makes it difficult to leave your car running while you open the garage door.

The increasing controls from the states, and changes in the administrative structures within institutions, have often led to statewide planning for computing. Changes in technology which have enabled networking have accelerated this trend. Plans have evolved which have been developed in concert between the ADP personnel within the institutions, institutional administration and state administrators. In some cases, implementation of plans has included state officials being designated as ADP coordinators or directors for the state.

This session is intended to create an atmosphere in which there could be external review of some of the individual state plans. State plans are presented by individuals who have been involved in the development of each, and critiqued by others not involved in that plan. We hope thus to provide external plan review while initiating some free exchange among the participants.

Educational Information Services, Inc.

by Gilbert G. Moser
EIS, Inc.

What is EIS? Educational Information Services, Inc. is a not-for-profit corporation in the State of New Jersey providing central computing resource to educational institutions. EIS was incorporated August 1, 1972. A number of the present EIS staff were with a predecessor corporation, New Jersey Educational Computing Center, Inc. ECC existed for just over two years and was funded primarily through a National Science Foundation grant with some matching monies from New Jersey. Over that two-year period, until August of 1972, ECC grew toward a shared higher educational computer resource. While some mistakes were inevitable, the experiment was deemed successful, and a decision was made early in 1972 that the shared resource concept should continue. Educational Information Services was incorporated to provide that resource.

EIS provides both a batch and interactive computing resource from an IBM 370/158 computing system which has been operational since October 1, 1973. Up to that time, during ECC days and through the first year of EIS, computing resources were offered from two centers: the Princeton University IBM 360/91 and the Rutgers University IBM 360/67. EIS has 20 high-speed RJE terminals coming into the system from various parts of New Jersey to provide local batch computing capability. The EIS 370/158 computer has 120 low-speed ports which are looking at approximately 180 terminals. Through them 6,000 validated ID's, principally students throughout the state, are using low-speed interactive services. Our principal interactive language, IBM's CALL/OS, supports BASIC, PL/1 and FORTRAN. In addition to CALL/OS, EIS offers APL, A Programming Language; ATS, which is a text-editing system, and COURSEWRITER III, which provides a Computer Assisted Instruction capability.

There are 48 people at EIS organized into five functional areas: computer operations; administrative systems and services; academic systems and services; corporate marketing; and community college services development. With final plans approved in March for the delivery of our new computer system in late July 1973, and with production scheduled for the fall 1973 academic semester, it would have been extremely difficult to build an operations capability within EIS. Consequently, EIS staff explored with Princeton University the feasibility of a partial facilities-management arrangement. At the present time Princeton, under contract with EIS, has responsibility for the operations of the IBM 370/158 system and for providing some systems support. On the EIS staff, three systems people are responsible for the interactive systems. Princeton has responsibility for the EIS batch systems. EIS has telecommunications and production control responsibility.

EIS is in the administrative services business. It supports administrative systems which are used by the state colleges and others. A student information system, CUSTOM, provides student registration, scheduling and grade reporting functions. EIS also offers a facilities-management system called EFIS. In addition to the CUSTOM and EFIS support activity, EIS has a staff of ten working on the implementation and modification of new financial and admissions systems. Much of what EIS has done in the past and continues to do is under contract, principally to the New Jersey State Department of Higher Education. EIS also acts as project manager for the state with commercial vendors who are building higher educational administrative systems.

A third group is involved in interfacing with academic customers. EIS has the traditional "circuit riders"; trainers and builders of documentation who communicate with the academic community. EIS' successful activity in this area over the past year has been demonstrated by the growth in use of academic computer applications.

EIS has two other functional areas: Corporate marketing is responsible for overall marketing coordination and customer services. When the customer has a problem, it is the responsibility of customer services to apply the appropriate EIS resource to solve that problem.

The final EIS function has responsibility for developing EIS service to New Jersey's 17 two-year colleges. EIS has not done a great deal of business in this area but expects to do so in the future. EIS also does some work in the secondary-school community in New Jersey even though a current priority is to service higher education. The secondary school market is a very promising one.

Educational Information Services' annual report for the year ending June 30, 1973, is available for those who may wish a copy. EIS is currently earning at an annual rate of $2.2 million. Approximately $1,200,000 is revenue from the sale of computing services with the remainder of income derived from conversion of major customers to the central system and from administrative systems development and implementation.

A ten-member board of directors, with constituancy representation roughly proportional to EIS' dollar business with the constituancy, directs the corporation. Rutgers, The State University, has two members on our board; the eight state colleges which are major customers have two seats on the board; and the independent schools, the county schools, the New Jersey College of Medicine and Dentistry and Newark College of Engineering, each have one member. The Board also includes a public representative and a member from the State Department of Higher Education.

EIS enters into contractual service relationships with all customers. Contracts are professional expressions of the services EIS will provide in return for payment. EIS works from a nearly exact entrepreneurial model with all of the plusses and problems that are inherent to any business. Because it operates on user fees, EIS must be excruciatingly service oriented; we must do what we promise to do.

EIS provides a necessary educational resource to New Jersey students and their faculties. The rapidly growing use of this resource is a measure of the corporation's success.

Critique of the New Jersey Plan

by Freeman Holmer
Oregon Department of Higher Education

It is one of the virtues of a Federal system that there is an opportunity for a variety of experiments without the risks associated with national commitment to a single approach to a felt need. The fact that EDUCOM is devoting workshop sessions to the subject of "Statewide Plans" is evidence both of a widespread belief that there is virtue in interinstitutional planning and of a variety of responses to that belief.

In creating Educational Information Services, Inc., the state institutions of higher education and the New Jersey Department of Higher Education have made a commitment to long-range planning to provide a wide range of computer-related services. The Educational Information Services (EIS) approach is one of several options. It is characterized by three significant characteristics:

- Management by a separate public corporation.
- Acceptance of a competitive role.
- Limited commitment to central systems development.

MANAGEMENT BY A SEPARATE PUBLIC CORPORATION

It is imperative to the success of a statewide computing consortium (or, indeed, to any consortium) that the members feel that their individual interests are fully protected. Establishment of a separate corporation is one means of providing a degree of assurance that this is the case. The real test comes, however, in institutional evaluation of the responsiveness of the corporation to institution needs. The dominance of Rutgers University in the corporate structure could be an eventual source of concern, but New Jersey's second assumption provides a safety-valve.

THE COMPETITIVE ROLE

EIS is not the sole provider of computing services to public institutions of higher education in New Jersey. In addition to its own computing resources, EIS serves as a broker of surplus computer resources at Rutgers University and Princeton University. This multiplies the options available to customer institutions.

However, while EIS serves as a clearinghouse for planning as well as a purveyor of services, the option for institutions to seek other computer-related services poses a hazard to the ultimate viability of the corporation. It is possible that the New Jersey State Department of Higher Education or the legislature may enforce economic sanctions, limiting institutional autonomy, but EIS is in no position to impose any sanctions whatever. The sanctions otherwise may run the other way.

The monopoly model is, of course, limited to public institutions. Such a model is less appropriate in northeastern states than in many other states because of the prevalence of private institutions. However, the case can be made for a monopoly role for centrally provided services to public institutions perhaps with optional participation by private institutions.

A *mixed* model even for public institutions may be the preferred mode. Such a mixed model might include central computing power and a substantial central systems and programming staff, while allowing a range of local decisionmaking and action in design, installation and operation of academic or administrative systems.

On the continuum from the competitive role performed by EIS to a complete monopoly (not known anywhere, yet), each state must identify the range of freedom in planning and operations that is appropriate to its circumstance. In the determination of systems to be designed, one can see a potentially uneconomic result of the EIS approach.

CENTRAL SYSTEMS DEVELOPMENT

EIS has a small administrative services and systems staff. It is also committed to the acquisition of externally developed systems and modules to be made available to EIS users. However, EIS does not have a fully effective mechanism or authority to avert the repetitive invention of the quill pen at each participating institution, except for the willingness of each institution to use established EIS packages. While the virtue of systems tailored to the unique characteristics of the using institution should not be minimized, redundancy in systems development can be a major source of waste in the use of limited manpower resources.

Central review of systems needs and collaborative determination of systems development priorities will, in many settings, result in more effective use of manpower and greater response to the universe of need. Such central software planning need not result in systems strait-jackets since common systems can include options that satisfy unique requirements.

The choices New Jersey has made in establishing EIS and the predecessor consortium constitute a major step forward in the provision of computer resources to the public institutions of higher education in New Jersey. EIS would be well-advised, in my judgment, to reconsider from time to time whether the corporate form, the competitive model, and substantially decentralized systems planning should be modified in the interest of a potentially more effective service to the institutions and the state.

Critique of the New Jersey Plan

by David McIntyre
Illinois Educational Consortium for Computer Services

Three questions can be addressed to those responsible for the New Jersey plan to gauge the effectiveness of the EIS organization. First, what is this plan doing to a very controlled economic system? Historically, the academic computer center has operated in a captive customer relationship. One asked for money from the state, the university received money and spent it on the inhouse computer. Occasionally some mavericks went out and bought time from timesharing services and commercial vendors. Under the New Jersey plan for computing for higher education the state is developing a third force that is neither an inhouse computer center nor a commercial organization. What assurances can one give to the inhouse organization that meddling with this heretofore closed economy will not do precipitous things to the inhouse capability? Extraordinary events can occur when one attempts to modify and control economic systems, at least on a national scale. Extrapolation would tend to make one very wary of what can happen to computing centers under similar circumstances.

Second, what is the formation of EIS doing to the morale and motivation of the people in academic computing centers? The people who surround the computer are often a university's most scarce resource, and are more sacred and more scarce than any machine. Think of what happens when colleges change computing equipment. It is done all the time. It is not fun thing, but one lives through it. However, think of what would happen if a college kept that equipment and changed all the people. There would be an absolute disaster. By far the people are the most important portion of the combination. The emergence of this third force can have a bad effect on the morale and motivation of the inhouse people. When the computer center director is faced with uncertainty, he passes that uncertainty down to the systems programmer who also passes it down. Costs of this nature are not yet apparent in New Jersey, but in Illinois uncertainty associated with centralization of computing services has cost between 10% and 20% of the effectiveness of inhouse computer people.

Third, where are the economies in EIS' plan? Where are the fundamental dynamic principles that indicate that this kind of specialized organization is more economical than the existing one? Typically, the argument is in the economy of scale. A central large machine can provide computing services to people who can't afford a big machine for a considerably smaller price per unit of computation. College computer center directors have to accept that argument because they've been making the same argument to their administrations for several years. However, if one really wants to realize economies of scale, one should start with an existing center, and spend money to increment the capability of that existing center thus moving further up the economy of scale graph to a point where prices per unit of computation delivered go down even further. In that manner cheap computation can be delivered to people who couldn't afford it before.

The final argument about the economics of centralized statewide computing must include counts of the number of people being paid to operate computers. A statewide plan cannot claim economy when the central organization employs a larger number of staff than would be necessary if centralized service had been built using an existing computer center.

Computer Services in the Oregon Department of Higher Education

by Michael A. Jennings
Oregon Department of Higher Education

For the past decade the Department of Higher Education has coordinated its computing by an Interinstitutional Committee on Computer Activities. This committee, as well as outside consultants, advised the Department to establish more explicit comprehensive planning for computing to insure effective utilization of all the Department's computing resources.

For the last two years, the Office of Administration of the Chancellor's Office has coordinated a cooperative planning approach to identify the most effective and productive ways of utilizing the computing resources of the Department of Higher Education.

Task forces looked at two major areas of higher education computing: academic computing including instruction and research; and administrative computing including student, personnel, financial, facility, library, and medical systems. A summary of the findings and needs found within the two major areas were:

Administrative Systems

- A lack of coordinated planning for administrative systems development in the Department.
- A redundancy of expenditures in systems development.
- Limited attention to compatibility of data definitions.
- Lack of standards for systems documentation and computer languages.

- In some cases, evidence of poor utilization of existing computer equipment for administrative data processing.

Academic Systems

- A lack of coordinated planning at the institution and statewide level for proper use of computing devices.
- An imablance among the institutions of computing resources for instruction and research.
- A scarcity of computing resources to support computer and system science curriculums at the university level.
- An inadequate computing resource and service mechanism to support present level of outside academic computing ($150,000 for 1971-1973).
- A potential redundancy of expenditures in the development of academic computing systems.

GOALS

In order to meet the needs of the institutions and correct the deficiencies, a Department planning effort established some assumptions and goals that provided a foundation for a comprehensive long-range plan for the Department of Higher Education.

- Provide a framework for data processing planning in *all* institutions.
- Provide a flexible and yet consistent budget plan utilizing, as possible, current resources.
- Identify a foundation for equitable for equitable data processing growth.
- Produce a reasonable developmental plan for computing with appropriate milestones of checks and balances.
- Cause minimum disruption of current personnel functions.

To meet the overall planning goals, a concept of a computing *consortium* among all nine of the public institutions was agreed upon. A *computer network* for the Department of Higher Education is planned. The overall goals of the computer network are:

- To provide equitable access for students and staff at *all* institutions of the Department of Higher Education to the computing resources of the Department.
- To maintain and enhance the quality of computing resources and services presently available to students and staff at all institutions of the Department.
- To maximize the efficiency of computing resources utilization in both hardware and software development.
- To strengthen the use of computers for administrative and academic computing.

To help effectively meet the overall goals of the network operational goals for each major computing area, administrative and academic, were identified. These operational goals were established to insure proper attention to the respective needs and functions of each area. They are:

Administrative Computing

- Establish a Departmentwide organization to coordinate administrative systems development.
- Develop and share statewide systems for common administrative needs.
- Standardize administrative data elements using National Center for Higher Education Management Systems (NCHEMS) standards and definitions as a foundation.
- Designate project leaders for each functional systems area on each campus.
- Enforce State Executive Department computer program language standards.
- Adhere, in general, to the following priority in the development/refinement of administrative systems priorities:
    - Student Systems
    - Financial Systems
    - Personnel Systems
    - Facility/Physical Plant Systems
    - Hospital Systems
    - Library Systems
- Where feasible, *buy* administrative data processing systems rather than develop them within the Department.
- Involve executives and high level administrators in administrative data processing planning.

Academic Computing

The goals for academic computing are similar to the administrative computing goals. This similarity is due to many of the common problems of the respective areas.

- Establish a Departmentwide organization to coordinate the service and development of academic systems.
- Develop and share statewide systems for common academic system needs.
- Designate project and discipline representatives on each campus.
- Involve executive and high level academic administrators in academic computing decisions.
- Provide an environment for all institutions in which computer-related disciplines can develop and advance.
- Conduct research in computer and information system sciences so that faculty can use computation directly in the teaching-learning environment.

ORGANIZATION

On September 26, 1972 the Board of Higher Education approved the proposed network plan stating that the responsibility for effective operation of the network rests with the Chancellor and the Vice Chancellor for Administration.

Computer network services will be provided under the policy guidance of a Computer Policy Council with a central staff and, in all instances, institutional staffs.

The Computer Policy Council, established pursuant to action of the Board of Higher Education at its meeting on July 24, 1972, consists of the Chancellor or his representative and a representative of each institution (appointed on recommendation of the institution executive). The Council is charged with policy-development in its broadest terms, including recommendations relating to network operations and to employment of operating personnel. It is understood, of course, that the actions of the Council are subject to the usual processes of budgetary review and to review by the Chancellor and the Board.

It is expected that the Computer Policy Council will be assisted in its activities by user task groups and subgroups. Some major groups are: the Administrative Systems Group; the Academic Systems Group; the Policy Planning (Budget) Group; and the Technical Advisory Group. It is understood that each institution and the Chancellor's Office require one or more full-time staff members to perform the following necessary services related to the operation of a computer network: scheduling and operation of terminal equipment; planning for and encouragement of computer services utilization; development of unique software systems; operation of local computer facilities. The central processing units (computers) of the computer network are operated by a staff responsible to the Computer Policy Council. For the foreseeable future, the Council contracts for the operation of the central network facility with Oregon State University.

The central staff performs a series of crucial functions, whether by contract with Oregon State University or otherwise. Figures 8.1 and 8.2 suggest the varieties of activities essential to operation of an effective network.

POLICIES AND RESPONSIBILITIES

In order to achieve the highest degree of success, the computer network requires the cooperation and support of all elements of the Department of Higher Education. Each institution must participate not only in the use of computing facilities, but must assist in the planning and the development of network policy and goals.

The computer network for the Department of Higher Education

Figure 8.1 Administrative Staff — Central Network Facility

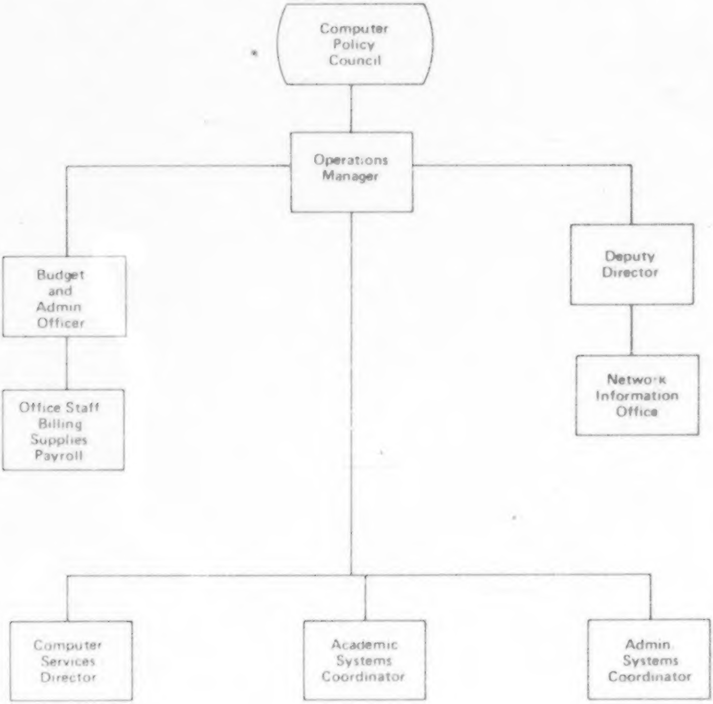
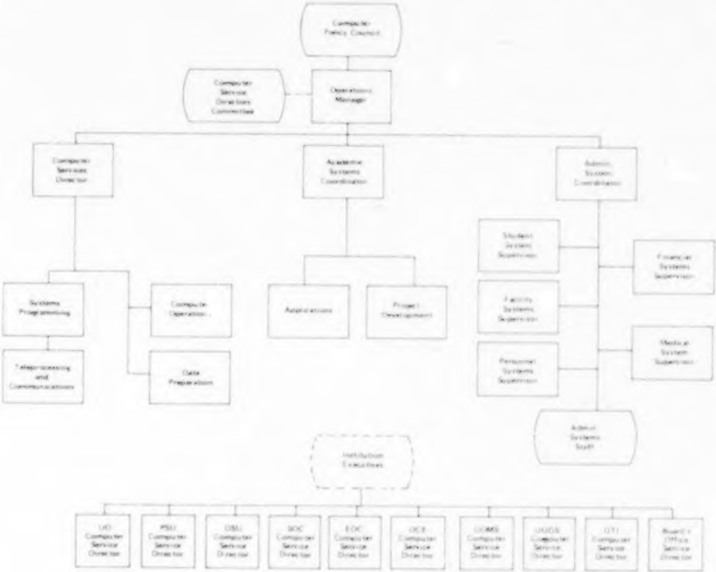

Computer Policy Council
Operations Manager
Budget and Admin Officer
Office Staff Billing Supplies Payroll
Deputy Director
Network Information Office
Computer Services Director
Academic Systems Coordinator
Admin. Systems Coordinator

Figure 8.2 Network Organization


Computer Policy Council
Computer Service Directors Committee
Operations Manager
Computer Services Director
Academic Systems Coordinator
Admin. System Coordinator
Systems Programming
Teleprocessing and Communications
Computer Operation
Data Preparation
Applications
Project Development
Student System Supervisor
Financial Systems Supervisor
Facility Systems Supervisor
Medical System Supervisor
Personnel Systems Supervisor
Admin. Systems Staff
Institution Executives
UO Computer Service Director
PSU Computer Service Director
OSU Computer Service Director
SOC Computer Service Director
EOC Computer Service Director
OCE Computer Service Director
UOMS Computer Service Director
UODS Computer Service Director
OTI Computer Service Director
Board's Office Service Director

includes centralized computing facilities, staff, and a variety of remote job entry (RJE) and interactive terminals. These facilities will provide each institution with a full range of computer services operated by institution staff. Ultimate computer network supervision and control will be the responsibility of the Chancellor's Office, with policy recommendations from the Computer Policy Council. Each institution will be responsible for scheduling and operating its terminal equipment, for providing staff and student assistance in the use of computer services, for developing unique software systems, and for operating local computer facilities.

The Chancellor's Office is responsible for the policies, plans, budgets, and programs relating to data processing systems development and implementation, for approving the acquisition and use of automatic data processing equipment, and for the computer network operations.

The Computer Policy Council's chief responsibility is to recommend to the Chancellor policies relating to the coordinated design and method of operation of the computer network.

Each campus maintains its own computer center, computer services director, or other appropriate service organization to coordinate institution computing requirements. In this sense the network for higher education operates in a decentralized fashion with each institution responsible for meeting its needs in instruction, research, and administrative computing using the network facilities. Each campus is principally responsible for providing staff assistance, production output, programming support, and, in cooperation with the network, assigning job numbers, saved file space and computer terminals. Each institution is responsible for administration of its data processing budget.

Oregon State University has the operational responsibility for the computer network for the Department of Higher Education.

In fulfilling this responsibility, the University: 1) manages all aspects of the day-to-day operations of the central network facilities including telecommunications; 2) administers the central network budget; 3) gives periodic reports and makes recommendations to the Computer Policy Council and Chancellor's Office concerning current operations and future network requirements; and, 4) undertakes special tasks and assignments requested by the Chancellor's Office or the Computer Policy Council.

Oregon State University staffs and maintains a Network Information Office to provide continuing assistance to each institution. The functions of this office includes training, consulting services, publication of a network newsletter, and the coordination and distribution of network manuals and written materials.

Training programs include curricula for central network personnel and institution computer center staff and selected short courses for instruction, research, and administrative users. These programs operate on a continuing basis with classes held both at Oregon State University and at each of the network campuses.

ADMINISTRATIVE SYSTEMS VIA THE NETWORK

One of the goals of the network is to implement common software systems to meet common administrative needs.

The Computer Policy Council assisted by the Administrative Systems Users Group reviews administrative system requirements and recommends shared statewide systems and standardized administrative data elements in several areas: Student Systems, Financial Systems, Personnel Systems, Facilities/Physical Plant Systems, Hospital and Clinical Record Systems, and Library Systems.

The central network staff coordinate the development and implementation of administrative systems which are implemented under guidelines and priorities recommended by the Computer Policy Council.

After working with network staff to implement systems, each institution is responsible for the maintenance of its data bases, scheduling the running times of administrative programs and production output.

ACADEMIC COMPUTING VIA THE NETWORK

Each institution is responsible for meeting its instructional and research computing requirements. The institution computer service center is the principal focal point to: 1) provide consulting and programming assistance in the use of any of the network facilities; 2) assign job numbers; 3) schedule the priorities and usage of remote terminals; 4) develop or adapt programs to meet the particular needs of the individual campus; and 5) assure that computing services are available on an equitable basis to all campus users.

The central network staff coordinates the implementation of computer packages that meet common requirements and will be widely used on the network, like simulation languages, statistical subroutines, instructional programs of proven utility, and scientific subroutines which support research computing applications.

IMPLEMENTATION AND EVENT SCHEDULE

Implementation of the computer network commenced on July 1, 1973. Because the project is a complex one, strict task-oriented planning was required including an allotted budget framework with proper check points for progress review. The plan was categorized in four major areas of implementation: hardware systems, software systems, personnel and training, and review and evaluation. Initial implementation has followed the plan.

The plan for network use of hardware systems focused on development of a central network facility at Corvalis between 1973 and 1975 through upgrading of the OSU-owned CDC 3300 with a CDC 3500 for academic computing and installation of a larger successor to the Controller-leased IBM 360/40 for administrative and academic computing. The PDP-10 and IBM 360/50 at the University of Oregon would be retained and RJE terminals would be installed on a quarterly schedule at participating institutions: Portland State University, Southern Oregon College, Eastern Oregon College, Oregon College of Education, University of Oregon Medical School, Oregon Technical Institute, University of Oregon Dental School, and University of Oregon.

Subject to further review by the Board of Higher Education, several leased computers would be removed beginning with the UO-leased IBM 360/20 by June 30, 1973. Figure 8.3 illustrates the schedule of installations and removals proposed. Hardware decisions are subject not only to further Board review and action but to refinement and modification in the light of competitive bidding. The implementation of these decisions will permit time-sharing operation (serving instruction and research users) remote-batch processing, and batch processing (particularly required by administrative users).

The objectives of software planning are to provide equitable access to the needed variety of programming languages, packaged tools, and coordinated systems development. Figure 8.4 gives a summary of implementation start times for different package and services at specific institutions.

Pursuing the recommendations and goals of administrative systems development that data elements should be standardized where feasible and administrative systems software should be purchased rather than developed, a nationwide search was undertaken to find a vendor experienced and successful in implementing student information systems. Systems and Computer Technology, Inc., (SCT) was selected. During 1971-73, SCT has developed student information systems specifications for the Department at Portland State University, the University of Oregon and Southern Oregon College, and has implemented an admission system at the University of Oregon. As a result of these efforts, SCT's ability to develop an integrated Departmentwide system which retains appropriate institutional differences in the operational areas of student information systems seems assured.

For a fixed price of approximately $400,000 SCT has agreed to develop and implement student information systems for participating institutions in the areas of:

- Admissions
- Student Information Systems Retrieval/Receiving
- Student Data Base
- Grade Reporting
- Financial Aid
- Course Request Processing

Figure 8.3 Network Hardware Systems Schedule

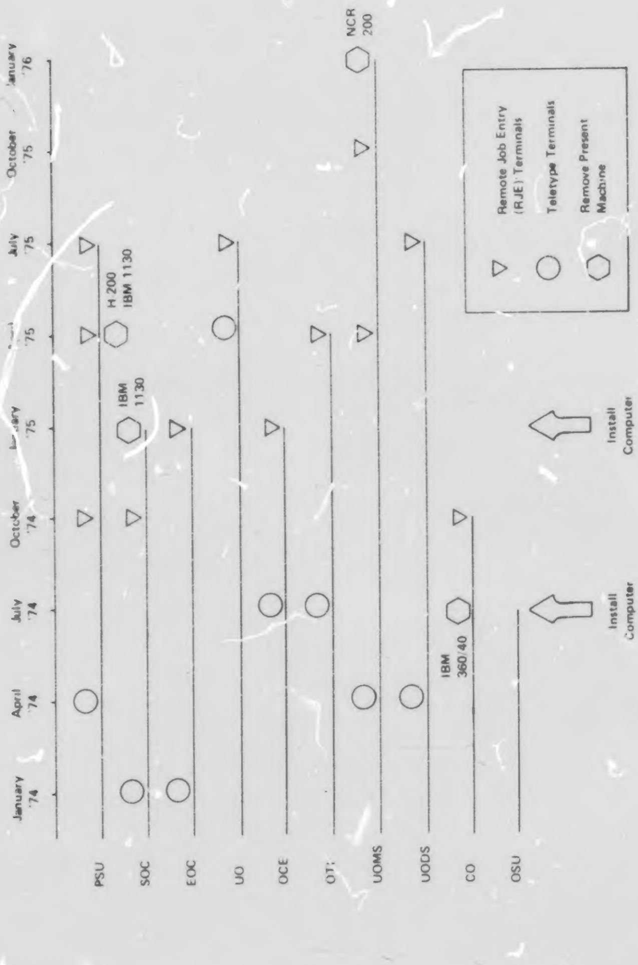
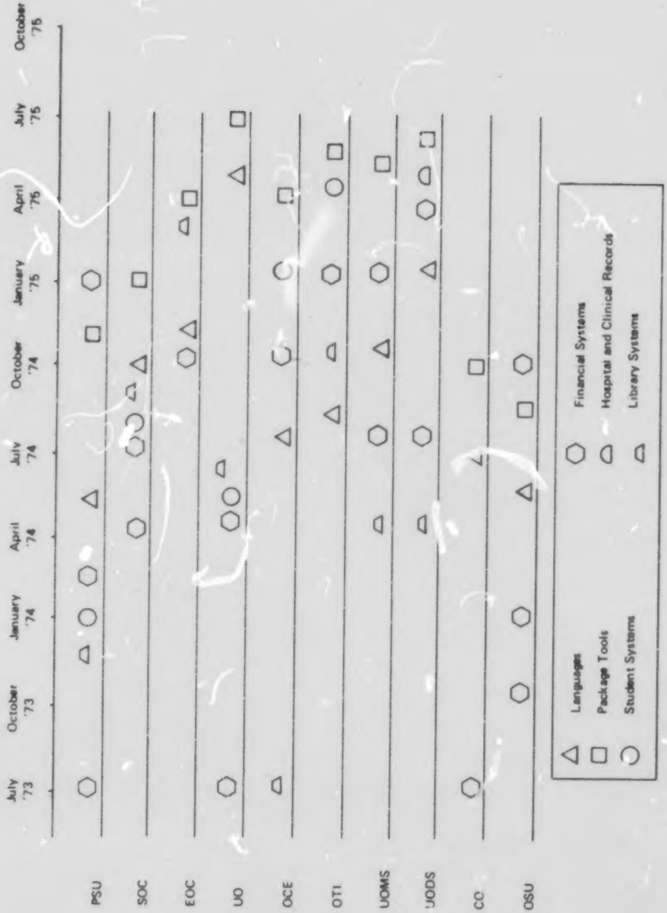


Figure 8.4 Software Systems Implementation Start Times



- Course Schedule Maintenance
- Student Scheduling
- Add/Drop
- Student Billing
- Alumni/Development Fund Raising.

Six institutions would participate initially: PSU, UO, OCE, SOC, OTI, and EOC. Additional institutions would be eligible for installation without license fee after the six institution installations have been completed. Development and installation will include: specification assistance and service; software (for a license fee); installation including debugging, testing, and so on; training of computer and student area clerical staffs; initial operation of each module or system for the first school term installed; complete detailed documentation of computer systems and SIS procedures; and provision of an Oregon resident SCT staff during implementation.

When the SCT contract has been completed, the resulting standardization of all data elements in student system areas will provide compatibility among all participating institutions for student information reporting requirements, both state and federal. All systems will be fully compatible with the NCHEMS data definitions and NCHEMS models and other procedures.

Because the administrative systems project will be implemented on a large machine resource of the central network facility, all institutions will have direct communication by RJE to their respective student systems for operation and file maintenance functions.

Another application now planned is a central library ordering and accounting system based on the Library On-Line Information and Text Access (LOLITA) system now operational at OSU and OCE. LOLITA can, when extended to all institutions, provide improved accounting services while creating a valuable central file of purchases. This system will share the same communication lines, and, in some cases, terminals as will be used for other network computer applications. While implementation of LOLITA is not dependent on the network, sharing of facilities and communication lines should appreciably improve the cost effectiveness of the LOLITA system.

Administrative systems at the UO Dental School are next in priority. Development is needed to improve the quality, accuracy, availability and utility of: clinical records, income and supplies accounting, and student registration and scheduling. To do this a common data base must be established for all departments and divisions within the Dental School and a flexible data processing and retrieval system must be installed to accommodate the Dental School's present and future needs as they relate to curriculum, student-patient progress monitoring, research efforts and administrative information demands.

As a means of solving some of the data processing problems which are unique to the Dental School, a program improvement allocation has been requested to permit the Department of Higher Education entering into a contract with an established outside firm to obtain: previously tried and tested computer programs designed specifically to meet the unique needs of dental schools; and the professional help and guidance necessary for the implementation of these programs on the network computing facilities of the Department.

Other areas of administrative systems development are under investigation which are consistent with software development priorities, and will be considered for action, after formal presentations have been made.

Based on a goal of a minimum disruption of people, the implementation of the personnel and training portion of the computer network plan calls for ten new positions in various data processing classifications. Most new positions are at institutions that do not presently have data processing personnel. The remaining data processing personnel positions of the Department retain 1973 levels with some reorganization but without relocation.

Personnel action begins with the consolidation of the data processing staffs at the OSU computer facility and the Controller's Office to establish the central network staff. Each institution not presently staffed for the network data processing plan will acquire staff three months before first equipment delivery. During these three months, the staffs of the respective institutions will receive training from the central network staff and appropriate suppliers of the equipment to be installed. Figure 8.5 illustrates the complete personnel and training schedule planned through 1975.

In order to maintain an appropriate evaluation of progress the Chancellor's Office will continually review progress as well as coordinate planning using the monitoring and change methods of the Program Evaluation and Review Technique (PERT). An evaluation team of institution computer service directors has also been designated that will continue to meet with central network management to assist in the implementation and review process. The Computer Policy Council will meet on frequent occasions to review the overall progress of development and implementation and to ensure proper policy guidelines to meet the needs of the participating users and institutions. Periodic reports will also be made to the Data Systems Division, Executive Department, to indicate the progress of the network and to coordinate activities with statewide requirements. Figure 8.6 illustrates the schedule for all review and evaluation activites up through 1975.

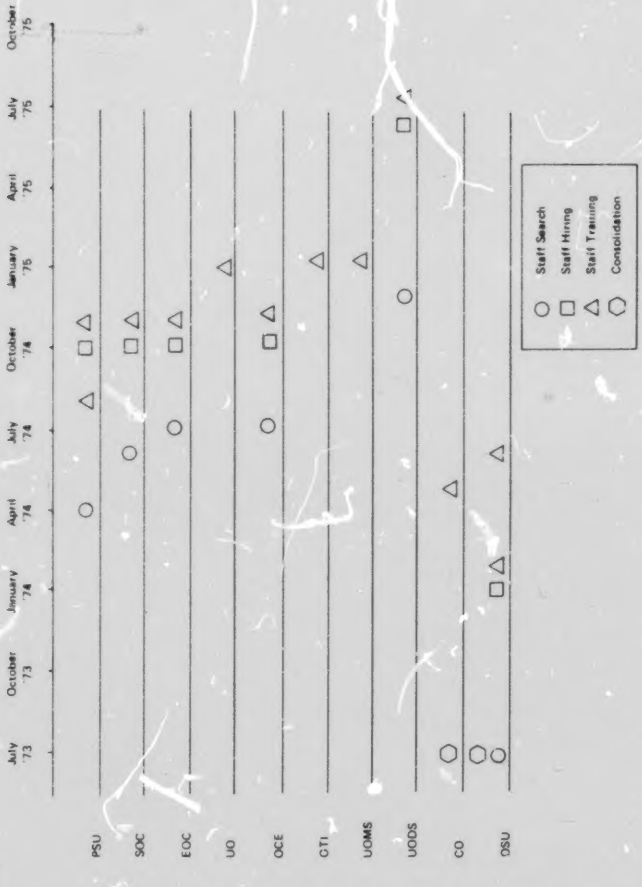
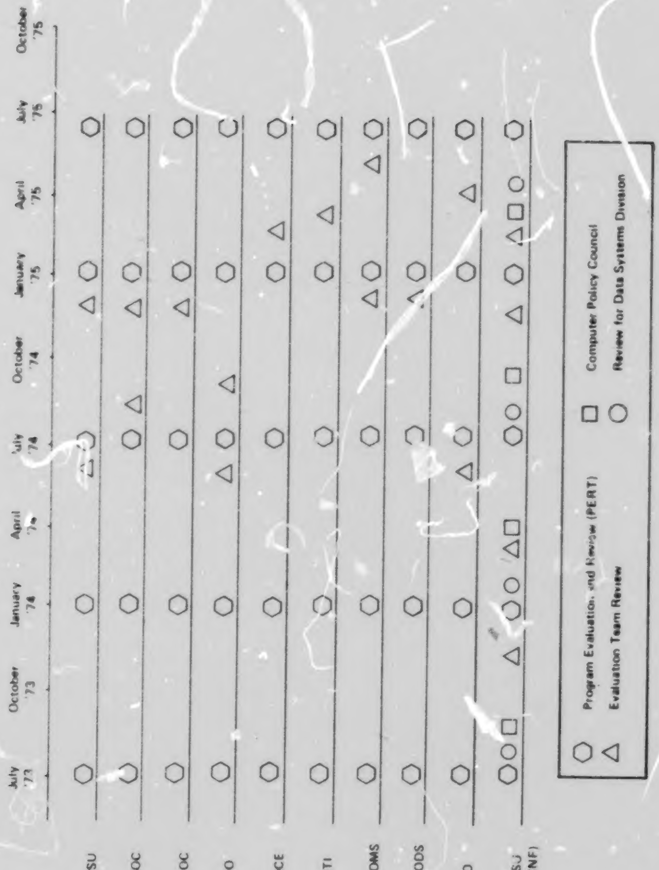


Figure 8.5 Personnel and Training Schedule

Figure 8.6 Review and Evaluation Schedule



FINANCIAL PLAN

One of the planning goals of seeking an alternative solution for higher education data processing was to utilize, as much as possible, the current level of data processing resources. The least costly solution over a five-year period for the data processing needs of the Department of Higher Education is represented by the proposed network. The resulting program improvement request represents an increase in the State budget for higher eduation computing activities – approximately 12.7% for 1973-75, approximately 9.5% for 1975-77, and less than 7% projected for 1977-79. These figures are compared with an average of 15-20% requested program improvement for each of the same periods without a coordinated and shared network concept.

Figure 8.7 is a summary of the financial plan related to the proposed improvements and the budget request of $888,326 for 1973-75.

Figure 8.7 Computer Services Summary Budget

FTE Employees are both academic appointees and
State Civil Service Employees

	FTE Employees		Budget	
1971-1973	173.7		$6,977,663	
1973-1975	180.1		$8,268,443	
Difference				$1,290,780
Base Budget Adjustment (5.7%)			$402,454	
Program Improvement				
(1) Network		$483,832		
(2) Student Information Systems		285,000		
(3) Centralized Library Accounting		119,494		
			$888,326	
Total				$1,290,780

The recommended Department data processing budget provides a financial base on which the network can develop.

Institutions will be responsible for their respective portions of the total network budget. Any further staffing and technical improvements at the institutions under the network concept will be reviewed and approved by the Computer Policy Council and Chancellor's Office.

The central network facility budget will be the consolidation of the respective budgets of the Oregon State University computer facility and the Controller's data processing facility. This consolidation holds the potential of staff cost savings after effective network operations have been established.

In order to effect orderly use and growth in computer usage for user institutions of the network, computer services, both academic and administrative computing, will be billed with charges based initially on volume of usage and a flat service fee. For all academic computing, instruction and research, a billing algorithm reflecting usage on a per job basis will be made which follows a traditional computer billing concept of central processing unit cost per hour for each job run. For administrative computing a flat service fee will be charged based on the administrative application (student systems, financial systems, library systems, etc.) and including other appropriate parameters. Since each student record contains approximately the same information and consequently the same computer processing and storage requirements a flat service charge on a per student basis would be appropriate.

Institution resources presently devoted to computing requirements are expected to continue to be used for the purchase of computing services and systems development either locally or through the central network facility. Institutional and central facility plans will be reviewed and coordinated through the Computer Policy Council and the Chancellor's Office to assure equitable access to the computing resources. Since there are some systems which should be developed centrally, great care will be required to assure an equitable allocation of the funds to be allocated to the central facility and the funds allocated to the institutions for specific institution requirements.

Critique of Oregon Network Plan

by Eugene Keenoy
Kean College

Comments in this paper reflect the opinions and experiences of a remote network user. Kean College, formerly Newark State College, is one of eight state colleges in New Jersey and has about 5,500 to 6,000 full-time students and about 7,000 part-time students with an FTE count of approximately 8,000. Kean College operates with an IBM 1130 computer with a 350 line per minute printer, allocating just over 50% of prime computer time to the academics and utilizing the remaining time as an RJE station for administrative work.

The spectrum of network framework goes from one extreme where the network is completely Chancellor-controlled such as Oregon, which has certain advantages, to an environment in which they form and contribute to the management of their own network. The latter framework is one in which the user community exercises complete governance over the policies of the network, one in which they satisfy their needs to the best of their responsibility within budget restraints. The New Jersey system probably falls somewhere between these two extremes. Users do have some degree of governance through representation on the Board of Directors. The Oregon plan, however, is completely controlled by the Chancellor. One has to be skeptical that the Chancellor's staff managing that network could be as sensitive to the needs of the user community as it should be. It was pointed out that there is a Policy Committee. The question before the house is "What clout does that Policy Committee have?"

In the initial stages of network formation Oregon State University will take on additional load as the host computer. The comment relative to "equal access" was interesting. Equal access is one of the most important characteristics for a network to maintain. Judging from three years' experience with an RJE work station, it is obvious that, without rigorous policies regarding equal access, it will be very difficult for Oregon State to not give priority to Oregon State jobs over the other users. If those contemplating centralization or state networks take away from the remote site the ability to do work locally and set priorities, then the least one can do for those remote sites is to make them all equal and keep any and all kinds of priorities to a minimum. The management and operation of the host computer by the state university cannot provide the same sensitivity to the user that a computer operated by a company such as EIS can provide. In a sense the term "equal access" applies to sensitivity as well as to job processing. User community requirements should carry the same influence with the system's programming staff at the network installation as the requirements of the state university. Judging from experience in New Jersey this is not always so.

Another difficulty which will probably be encountered by the Chancellor's staff operating the Oregon network is maintaining the quality of personnel. The private company approach to network management, such as New Jersey's EIS, has a better chance of success in this area because it does not have to adhere to the same salary guidelines as a state civil service system, can provide more incentives to their employees, has more freedom to hire and fire when necessary, and is not burdened by the red tape inherent in the state bureaucratic processes.

One of the requirements for a successful network that is missing from the Oregon presentation was a plan for the education of the users. The term "users" includes everyone at the remote site from the person operating the work station to the president of that college. The education of the users at the operational level will of necessity be technical; it will include instruction in the Job Control Language, the HASP commands, OS Utilities and file management concepts. At a higher level there is a necessity for education concerning what a user can reasonably expect from a network. Such training should define responsibilities, should make a commitment to turnaround time and explain what turnaround time means to demands of administrators for information, and should demonstrate to administrators that the price paid for the turnaround time delays is indeed less expensive than the total cost of running autonomous local installations. One of the biggest problems in managing a remote installation is the lack of understanding among "customers" of why it takes so long to process jobs. They have no understanding of priorities, no feeling for what the effect of set-ups are on job turnaround and no concept that there may indeed be more than one hundred jobs awaiting execution at the time their job is submitted. The network management or sponsor can do a great service to the user community if he can give college administrators and their staff a basic understanding of what network processing is all about. While it is true that this education could be done at the local level by the local director of computer services, the explanations tend to seem more like excuses than basic premises.

While it may be premature in the present planning stages, documented performance criteria for the network ought to be included in the Oregon Network Plan. How do they propose to evaluate whether the network is performing satisfactorily? The product that the network has to sell is service. There should be a mutually agreed upon set of turnaround criteria as an acceptable standard for that service which must include the variables of CPU time, core size and number of set-ups. In New Jersey institutions have agreed upon a set of turnaround criteria with EIS with the expectation that EIS will meet these criteria 80% of the time over any one month. There are financial penalties associated with failure to meet these criteria.

One of the advantages to the Chancellor-Controlled Oregon System is the ability to achieve a greater degree of standardization in data usage and coding than is possible with a group of colleges acting independently without regard for the coordination of how data elements are coded or how programs or processes are documented. Without such standardization almost identical programs are written by each college solely because each has coded data elements differently. This means more program libraries take up disc space and program maintenance costs are exaggerated.

The Oregon proposal to provide fixed pricing for administrative processing ought to be reconsidered. Glaring differences are apparent on the manner in which eight New Jersey state colleges handle administrative processing on the network. Some are very efficient and others are very inefficient. A fixed price scheme which forces the efficient user to absorb part of the costs of an inefficient user is inacceptable.

One issue in network economics is monopoly and related incentive. I contend that the private company (EIS) approach to networks provides more incentive for provision of user services than does the Chancellor-controlled type of network. The private company has more flexibility in salary ranges and ways to reward people for proficiency or capability than does the bureaucratic system. The incentive of EIS as a relatively small company in the service bureau business trying to make a go of it is pure and simple, a stake in their own future. While for all practical purposes the company may be a monopoly, it still cannot survive with a dissatisfied user community. It is incumbent upon the company to aggressively market their services and achieve the necessary revenues to support the systems and staff on hand. While an organization like EIS is a "not-for-profit" company it is also a "not-for-big-loss" company. Further, although technically "independent," the chancellor's office has more than a little to say about EIS policies and future planning. Experience has also shown that, where the user colleges were very dissatisfied with some aspect of EIS operation, the president of the college contacted the chancellor in addition to making the dissatisfaction known to EIS. EIS then experienced pressure to perform from two sides. While the extent of the Chancellor's office control over the local network may not be appropriate, one has to admit that two-sided pressure has to keep network management and staff on their toes.

The Oregon proposal in comparison to the New Jersey environment tends to make the network less competitive. In New Jersey a billing algorithm charges on-line storage at two cents per track per day. With numerous applications in which on-line storage would be advantageous, the user colleges calculated what it would cost to maintain files on on-line storage and reviewed those costs relative to each data processing budget. EIS will probably have a lot of unused on-line storage on hand for a long time. EIS staff have already been told that, if they expect users to take advantage of this capability, they are going to have to sharpen their pencils in that portion of the billing algorithm. The network will have to weigh the advantages of increased user activity against a lower charge for on-line storage. In the Oregon environment in which the state controls not only the funding but also the network, a different approach might be taken toward resolving such cost problems.

Oregon and other states contemplating networks or even maintaining existing networks ought also to consider career paths for the employees at the remote sites. At Kean College with a staff of three very competent programmers one must ask "Where are these people going?". If the state decides to contract for a large student information system from an outside vendor, acquires a financial system from Florida, and gets a library system from someone else for all eight state colleges, when does the Kean College programming staff get an opportunity to work on any kind of challenging programs? How do they become expert in the use of COBOL or PL/I? How do they grow professionally? There is a lot of talent out there in the boondocks. The next time a challenging job comes up, why not have some of these people participate in its implementation? If the network maintained a central staff of highly competent experienced people, and the programmers from the remote sites were given the proper experience, then the career-path for the programmers might be to eventually assume one of the more responsible positions on the network. At least added experience would qualify local programmers for more advanced titles and salaries locally, even if they choose not to go to the network. With the present arrangement it is inevitable that quality programming staff will leave the local sites. Any talented programmer will also have the ambition to move out to an environment where he or she can participate in more advanced projects. This problem should be addressed in the proper perspective with all of the other network and centralized problems.

The most significant comment one could make on the Oregon plan is that the key to the success of the Oregon Network is whether the policy committee has enough authority, and whether it truly represents the user community.

The Florida Regional Computing Center Plan

by James Morgan
Florida State University System

Statewide plans for computing in Florida are exactly geared to the environment there. In a different environment, state and university officials might have done many things differently. The State of Florida has a rather odd shape and has public universities studded all around the edge of it. In West Florida the University of West Florida in Pensacola with around 3,000 to 3,500 students offers third and fourth year program leading to the Bachelors degree and some Master's Degree programs. In the northern part of the state at Tallahassee is Florida State University which is a broad spectrum institution that covers baccalaureate and advanced graduate programs and Florida A & M University which was the predominantly black university in Florida. A new university in Jacksonville, the University of North Florida, enrolls approximately 2,000 students in a third and fourth year and beginning graduate program. Florida State University in Tallahassee and the University of Florida in Gainesville, Florida enrolls students from the freshman to advanced graduate level. At the University of South Florida in Tampa baccalaureate and graduate programs are offered but on a much smaller scale from that at the University of Florida and Florida State. Florida Technological University in Orlando which enrolls freshman through beginning graduate level students was built there and called technological because of the concentration of aerospace industries in the Orlando area. However, the establishment of Disneyworld, near Orlando has substantially changed the character of the area and caused some problems in adjustment for the University.

In the Southern part of the state, Florida Atlantic University in Boca Raton enrolls approximately 7,000 students in its third and fourth year baccalaureate and beginning graduate program. Having been established in 1963, Florida Atlantic University is an older institution in Florida. Florida International University in Miami is a new university which last year opened with 4,500 students enrolled in third-fourth year and beginning graduate programs. Enrollment has increased substantially in 1973.

Because the State University System in Florida must serve all of these institutions, planners must talk about regional computing in the State of Florida. Before any regional system was considered substantial computer power existed on many campuses. At University of West Florida an IBM 360/40 with 64K of core did everything that they needed to have done. At Florida State University a CDC 6400 handled computing for instruction and research and a Honeywell 1200 handled computing for administration. Florida A & M University had a 1401 with 12K. The University of North Florida and Florida International were not open at the time. For instruction and research the University of Florida had an IBM 360/65 in addition to an IBM 360/50 for administrative support and IBM 360/30 in the teaching hospital. The University of South Florida had an IBM 360/65. Florida Technological University had a Honeywell 1200, and Florida Atlantic had an IBM 360/40.

When regional service was first considered, some problems had to be solved fairly rapidly. First, one could see increasing pressure to provide the full spectrum of computer services on all campuses, and it was obvious that this would not be possible with the existing hardware on some of them. Second all campuses exhibited a fantastic reluctance to enter into any kind of computing arrangement with another university where they had no voice in the operation of that computer activity.

At the time of the original plan, there were less than 100 interactive terminals in the whole state and no remote job entry terminals. These terminals were concentrated with approximately 80 at the University of Florida and the remaining number at Florida State University.

The campus computer centers were marginally staffed, and there had been very little opportunity for exchange of systems and programs between the institutions. Although exchange of systems might have been advantageous in some cases, there was no opportunity at all in others. For example, there was nothing one could have moved to the IBM 1401 at Florida A & M which could have solved the administrative computing problems there.

Within the State of Florida the EDP Division of the Department of General Services is by law charged with total responsibility for the acquisition and operation of computers. If staff at public colleges or universities want to lease, purchase or otherwise acquire a keypunch, one must go through this state agency. If the cost of the purchase were high, the decision would have to be considered by the cabinet which sits as the EDP Control Board in Florida. Significant kinds of problems with computing in higher education in Florida were not unique. As the EDP Control Board of the GSA began to exercise more of its powers of review, representatives of the colleges and universities found it hard to justify on a rational basis some of the activities then going on in computing. For example, how does one defend nine separate payroll systems? How does one defend acquisition of an IBM 360/65 with a 256K core for a single institution? It was necessary for the State University to develop a plan to regain the initiative in computer planning for the university system.

Basically, the State University System plan got the attention of all the universities in the system. The plan advocated using a smaller number of larger computers, to take advantage of economies of scale in computer hardware. Economy was not stressed because, along with the economy that came from the larger hardware, State university officials envisioned a vast and substantial expansion of computer services in support of instruction. To take advantages of economies of scale in hardware, to eliminate the duplication of systems effort, and to expand instructional support, the initial plan recommended the creation of: four regional centers for administrative processing; one central center for instruction and research support for the whole system; and consolidation of systems staff. In the recommended functional systems arrangement, a single system staff would support statewide systems for student records, administration and other functional areas.

REGIONAL CENTERS

At the time the initial plan was released the University in Miami was still without computer hardware and was dependent upon the IBM 360/40 then at Florida Atlantic University. The administrative and academic vice presidents and the computer center directors of the two institutions met and began to consider how they could effect the kind of regionalization of computing services and hardware sought by the plan without closing the universities in the process. Within two months these institutions submitted a plan for a regional center (location unspecified) which would provide the complete spectrum of computing support for both institutions. After approval by the Board of Regents, this plan became the implementation plan for the first Florida regional center.

In this particular region, the regional computer center supports administration, and instruction with only unusual kinds of computing service provided by either Florida State University or the University of Florida. For the first time all data processing personnel in that region belonged to the regional center rather than to the participating universities. This worked out well because people were strongly in support of the idea and realized that they were getting some things from the regional center that they would have been totally unable to get on an institution-by-institution basis.

In the fall of 1971, a regional center plan was adopted first by the Florida Board of Regents and second by the State Cabinet which gave the framework within which other regional centers could be developed. The plan established that four regional data centers under the control of the institutions would operate central hardware with remote job entry type devices. Each one of our regional centers is under the control of a Regional Policy Board. The policy boards, in most cases, consist of the academic and administrative vice president of each one of the institutions involved, and a fifth or seventh member from the chancellor's office. Since one member of each Policy Board represents the Chancellor's office, that office can coordinate activities of the regional centers while each center is still controlled by the institutions.

For six months following adoption of the one page plan for regional centers, the groups of institutions involved with each center worked together to produce an implementing plan. The regional centers began to become operational in June or July of 1972 and some of them have now been operational 16 to 18 months.

Wherever possible, a regional center was built on existing hardware. If this were not possible, the regional center staff in conjunction with the EDP division of the state worked out a series of specifications for bid for a computer. In July of 1972 the regional center staff in Miami installed a UNIVAC 1106 at Florida International University and began to phase out the IBM 360/40 then located in the computing center in Boca Raton which became a terminal data center. The UNIVAC came in on schedule, the 360/40 was phased out on schedule, but all did not work beautifully. There were six traumatic months from the time that UNIVAC system came in until all systems were ready.

From June 1972 to May 1973 the University of South Florida upgraded an IBM 360/65 computer through the bid process from a 256K to 512K core, and replaced all disks with plug compatible units. During the year the center also phased out an IBM 1130 and a Honeywell 1200 replacing both with Data 100 type terminals to a regional computing center.

The regional data center in Tallahassee had more significant problems because of the lack of availability of administrative hardware in that particular region. When the regional center plan was initiated, Florida

State University in Tallahassee had a Honeywell 1200 and a CDC 6500. The first decision that the regional policy board had to make was whether to buy service for the CDC 6500 or to initiate separate procurement of an administrative computer for that area. The decision that they made was influenced to some extent by the availability of hardware. The North-West Regional Center was implemented on the IBM 360/40 which was phased out at Florida Atlantic University. When the 360/40 was moved in September of 1972 to the Regional Center at Tallahassee, everyone clearly understood that any production done in that year was pure bonus. The machine was there to allow a long period of time for conversion of the Honeywell 1200 at Tallahassee and the IBM 1401 at Florida A & M. In August 1973, the 360/40 was replaced by a 360/50 which had been purchased by the University of Florida, and the 360/50 became the regional computer for the Northwest Regional Center. Although the computer for the Regional Center is from IBM, the core, tapes, and disks are from Ampex, and the terminals are from Data 100. The simultaneous conversion at Florida A & M University from the IBM 1401 system to systems on the 360/40 and 360/50 consisted of looking around the system for operating student programs and financial programs to augment the basic Florida State University System accounting module. In September 1973, Florida A & M University began operation on a completely borrowed set of systems.

In Gainesville, the University of Florida Computer Research Center ordered an IBM 370/165 when the decision was made by the Regional Policy Board to put all applications available on that computer. In March 1973 the 360/30 in the teaching hospital was phased out and in July the 360/50 in the administrative center was moved to Tallahassee. Remote sites are without exception serviced by Data 100 type terminals with a variety of devices on them, including tape drives. Since everything else has changed, most of the institutions have eliminated keypunches and have gone to Infrarex key to disk system. The little tape drive on the Data 100 terminal gives one a good way to get data into the computer.

Since the Regional Center Plan has begun to be implemented the relationship between the Chancellor's Office and the EDP division has changed remarkably. At least some of the vulnerability of the State University System has been removed and relationships with the division are much more pleasant. In addition EDP Division personnel have been of substantial help to the State University System in 1973 in getting things done. With the processing procedures of the state government, one must recognize that great effort is required to go through the bid process to upgrade a 360/65 like the one in Tampa by adding core, disks, and communications control from other vendors. It takes a fantastic amount of cooperation to go through the bid process to order all these devices and the State University System got it.

Critique of Florida Computing Plan

by George Struble
University of Oregon

A regional center system like that of Florida is an intriguing idea, and is offers several advantages over systems that try to centralize everything. There are advantages in asynchronous conversion; when the next steps have to come along, one does not have to convert everything, everywhere, all at once. One can shift some of the workload from a computer to be phased out to another center. Personnel can concentrate on conversion of part of the computing workload at a time, even borrowing some staff from other regional centers. Another advantage is that regional centers can offer a richness and variety of instructional and research facilities. The richness and variety are more a hindrance than a help in administrative applications, but can be quite a help in academic applications, because the applications themselves are quite varied. Variety of hardware can contribute to the ability to serve a variety of academic needs, and variety and richness of software are even more helpful.

In the Florida plan, however, there appears to be some imbalance in planning for administration between the administrative and academic needs. The original proposal, by which the administration hit the institution over the head and got their attention, seemed to centralize service of academic needs in about two or three sentences. Although the proposal included a detailed staffing plan, that plan did not provide for any academic support personnel, either at the center or at the institutions! That specific problem has been remedied in subsequent planning, but one still gets the impression that the priorities have gone to administrative development rather than to academic development. Although both administrative and academic development is needed, the principal aim of higher education is the education of students. One wonders whether the network planners always kept that in mind.

In development of an information service, there are some advantages in developing regional centers in order to effectively serve the individual institutions. Dennis Fife, of the National Bureau of Standards, observed that university or government personnel do not even know yet how to construct an effective network information service. Especially for academic services which have to be provided to a transient public which always needs training and good access to information services, a responsive and accessible information service is crucial to an effective computing service. Having each center serve primarily only a few nearby institutions certainly reduces the scale of the problem.

It would be of interest to know whether there are plans to overlap the service areas and to share the facilities between the regions in Florida. Tieing the regional centers together, or tieing a terminal data center to a regional center other than its primary host would help provide greater richness of service, but it certainly raises problems too, problems of funding, loading and information service, especially if there is a lot of overlap of service.

Finally, Charles J. Mosmann has noted, in Chapter 8, the distinction between networks based on provision of good service and imposed networks. Despite the fact that the institutions in each region in Florida made proposals concerning the organization of each regional center, and have regional policy boards, it seems clear that the Florida network is an imposed network. One indication of how good the service eventually becomes, will be the number of institutions not mandated to be in the system which come into the system and buy service from the regional centers.

Computer Activities in Illinois

by Timothy Griffin
Illinois Department of Finance

A state agency such as the Department of Finance in Illinois has many problems in running a consolidated data center for all agencies reporting to the Governor. The difficulties lie in two primary areas: the initial consolidating of the individual agency's hardware; and making the data center responsive to the needs of its customers. In the initial consolidation, there are the obvious problems of invasion of rights of the agencies' data center managers who naturally want to hold on to their own equipment regardless of the economies. This is no small task. It took the State of Illinois 18 months to create a data center with 32 agencies as customers. The job is not complete in that 3 agencies still have independent data centers, but 90% of the task is complete. Making the central data center responsive to the valid and oftentimes demanding needs of its customers causes great difficulties. Right now, the major objective in Illinois is to create a stable environment within the data center. While major strides have been made in this area, the data center still does not have the complete confidence of all state agencies which are customers.

Problems encountered with the Illinois General Assembly are typical. Often the Assembly acts on insufficient or inaccurate information, sometimes leaving the data center with inadequate funds to provide central data processing services called for by statute.

In Illinois the public universities acted before the General Assembly voted to include them in a central data processing facility under a State agency such as the Department of Finance. The problems which would have been created by such centralization are obvious. Illinois public universities have created a private corporation called *The Illinois Consortium for Computer Services* which seeks to consolidate computer services among the state universities and colleges. ICCS is described by Dr. Brody in the following paper.

What is IECCS?

by Ronald Brady
University of Illinois

EVENTS LEADING TO THE FORMATION OF THE IECCS

In 1971 the executive branch and legislature of the state of Illinois endorsed "The Illinois Master Plan for Applying Computer Technology in the 1970's," called "Impact 70's." The plan called for the consolidation into a single organization of all computer activities in the state agencies which report to the governor. Legislation was adopted which provided the appropriate authorities and finances to accomplish this consolidation. Education was not considered a state agency which reported to the governor, and as such, was not included in the consolidation, but it was clear that higher education as an "industry" would be expected to act in a responsible and economic manner to avoid replicating capabilities, systems, and mainframes.

Shortly after the adoption of Impact 70's, a private consulting firm was employed by the Illinois Board of Higher Education (IBHE) to access and analyze the effectiveness with which the institutions were using their computer resources. This firm conducted an "audit" of most of the state universities. Their findings indicated significant inefficiencies in the way the institutions were developing administrative systems and distributing computer resources. Acting on this information, the IBHE dramatically reduced recommendations for support of computer personnel and hardware in the regular budget submitted to the Governor's office and the Legislation.

Concurrently, the IBHE and the Department of Finance assembled a large task force of more than 100 people representing community colleges, private colleges, and the public institutions in Illinois to study the problems of educational computation and make recommendations to the board. The task force's deliberations and recommendations were included in a report, adopted by the IBHE, called *The Statewide Plan for Computing Resources in Illinois.* This report called for two significant actions: formation of a public interest corporation to own and operate most of the computers in the higher education system; and consolidation of all administrative system development activities.

In response to verbal preliminary recommendations from the IBHE, the Joint Council On Education, composed of the presidents of the thirteen public institutions, formed two task forces to study different aspects of computing in higher education in Illinois. One task force reviewed the then current organization for provision of computing services in the colleges and universities. The second task force developed plans for unified computer support for public higher education in Illinois. As a result of the recommendations of the task force concerned with organization IECCS was formed in September 1972.

The Illinois Educational Consortium for Computer Services (IECCS) is a not-for-profit corporation formed to:

- advance the development and use of computing technology information systems in institutions of higher education as a means to improved management and more effective education.
- encourage, promote, plan, develop, and provide increasing cooperation, coordination and sharing between and among member institutions, and all institutions of higher education in the state of Illinois, in the utilization of computer equipment, facilities, systems, services and personnel in order to achieve improved cost efficiency and strengthen and enrich computational capabilities for the performance and support of the respective educational missions, functions and operations of the members.

In short IECCS has been designed to do good things for educational computation.

In Illinois the thirteen senior public colleges and universities are administered by four boards and constitute four systems of public higher education. The University of Illinois which includes the University of Illinois at Urbana-Champaign, University of Illinois at Chicago Circle, and University of Illinois Medical Center is governed by a university board of trustees. The Board of Regents oversees the operation of Illinois State University, Northern Illinois University, and Sangamon State University. A separate Board of Governors oversees Eastern Illinois University, Western

Illinois University, Chicago State University, Northeastern Illinois State University, and Governors State University. Finally, under a fourth board of trustees, Southern Illinois University includes Southern Illinois University, Carbondale, and Southern Illinois University, Edwardsville. These four systems or governing boards of higher educationa are the members and owners of the IECCS.

The IECCS is directed by a Board of Directors composed of representatives from each of the four systems and representatives of two cooperating groups, the Illinois Community College Board and the private institutions of higher education.[1]

PLANS AND ACTIVITIES

Long range and general plans of the corporation are:

- To act as the vehicle through which participants may jointly fund, jointly use, and jointly control resources, including software, hardware, and other products.
- To act as a forum through which the institutional planners have some information to estimate costs for implementing plans in some way other than entirely by themselves.
- To make available to the institutions the advantages of a corporate entity which can offer flexible funding arrangements and rapid response to implement decisions.

Immediate plans and activities include:

- Formation of the Mid-Illinois Computer Center to support five of the smaller schools located near the middle of the state.
- Securing of financing to acquire and lease back to the public schools new computer equipment.
- Procurement and development of a library support system.

Intermediate range plans for the IECCS will be greatly influenced by the actions of two other groups outside the purview of the corporation. The University of Illinois is currently planning for the deployment of its research and instructional computer facilities, and a unified statewide data communications effort may be developed by the executive branch of the state government.

REFERENCES

1. Dr. John E. Corbally, Jr., President, University of Illinois (Chairman); Dr. Ben Morton, Executive Director, Board of Governors (Vice Chairman); Dr. John T. Bernhard, President, Western Illinois University; Dr. Ronald W. Brady, Vice

President for Planning and Allocation (U of I); Dr. James Brophy, Academic Vice President, Illinois Institute of Technology (Non-Voting); Dr. Gene Budig, President, Illinois State University; Dr. David R. Derge, President, Southern Illinois University, Carbondale; Dr. Richard Fox, Associate Secretary, Illinois Junior College Board (Non-Voting); Dr. Franklin E. Matsler, Executive Director, Board of Regents; Mr. David Nyman, Associate Director, Illinois Board of Higher Education (Non-Voting); and Dr. John S. Rendleman, President, Southern Illinois University, Edwardsville; were members of the Board of Directors in September 1973.

Statewide Networks for Computing

by Charles J. Mosmann
University of California, Irvine

One can distinguish two kinds of computer networks in higher education: those that emerge to serve a more limited set of purposes and then attract more customers by means of high quality service and/or low price; and those that are imposed on a system or set of users by some agency responsible for providing resources. From the user's point of view, these kinds of networks are very different. Networks that succeed in a free environment must adapt themselves to the needs and interests of the users. When networks are enforced on the user, however, he or she has little alternative but to adapt himself or herself to the system. In this case, the system has little incentive to develop either to economy or responsiveness.

In today's environment of multiple alternatives, the on-campus administrator's principal job must be to understand the questions regarding computing resources in terms of *educational* realities. These questions are not primarily technical, or economic, but have to do with the value of computing as a service, among other competing services, for the instructional and research goals of an institution. The administrator must understand what he is buying, in terms of the quality of the instruction and research provided, when he makes computing decisions. He is then in a position to defend the needs of his institution. When he understands the needs, and the available alternatives, he will be able to select among them the one (or ones) that represent the least cost, the least commitment, and the least risk. He will be pleased to let others, like the UCLA system described by Kehl, take the risks of long term hardware commitments.

If an institution decides to develop a computing service and allow its on-campus computing facilities to become a resource available to other institutions, it is important that those making the decision understand why their institution should want to be in the position of providing management and taking the management risks of developing a resource for other institutions to use.

The EDUCOM study of Academic Computer Planning in the American States and Canadian Provinces presently underway, is investigating the role of multi-campus governing boards and state systems of higher education in creating and encouraging the development of computer networks or statewide systems for computing. This project, funded by the Exxon Education Foundation, is currently under way and only tentative conclusions can be presented. However, a few points for discussion are apparent.

First, there are a wide range of models or forms of systems although the influence of geography, economics, and the nature of the politics or management is quite great. Some computer networks imposed on state systems of higher education have created equality of access by limiting the freedom of the members. In order to make the computing resources available to all the students in all the institutions in the system, the innovators and leaders in computing have been constrained to conform to the same patterns of behavior as those with little or no experience in computing. In fact excellence has been given up in favor of economy and democracy. Insofar as this is a conscious decision based on the requirements of that system in that state, such a decision cannot be criticized. Other networks reflect the need to market services and are highly user oriented. Many networks reflect the political, financial, and instructional realities of the systems of education they serve and are bound to be successful in that environment. Some plans for networks currently in development, however, appear unrealistic and unlikely of ever coming into being.

In the context of these divergent efforts one can propose the following hypothesis. Computing is not so unusual a resource that it does not follow some rules imposed on the structure of an educational system by external forces. The centralization of computing including the development of networks or regional centers or "super-centers" is linked to the phenomenon of more centralized control of higher education generally. Over the past decade, more and more states have replaced coordinating boards of education with governing boards.

Control of the computing resource is closely linked to control of the functions computing supports. If administrative data processing functions are moved to the state or provincial capitol or the board of higher education, so ultimately will administrative decision making be moved.

Fear of this centralization is the basis for much of the resistence to centralized data processing on the part of campus administrators. If, further, institutions lose control of institutional academic resources, including computing, then they lose control of academic decision making as well. Those responsible for academic decision making on campus must exhibit the same concern as that evidenced by administrators when faced with centralization of computing resources. The decision to centralize computing, has more than technical and economic components; it effects how a state and system institutions are going to do education in that system.

PART III

MANAGEMENT
INFORMATION
SYSTEMS

119/120

Chapter 9

An Overview

by Harold Wakefield
State University of New York

It's obvious that interest in that very elusive MIS monster hasn't diminished and that the obvious problems in developing management information systems haven't changed. However, the major problems are gradually being solved and there is much more optimism evident in the development of management information systems now than one year ago.

Generally speaking, the panelists who are involved in serious MIS development have completed about 50% of the development work. However, with new demands being continuously added this work may never be 100% completed. At the State University of New York over two years of effort on a continuing basis has been employed in the development of uniform information systems. It is obvious that many colleges have subsystems of an MIS that come close to satisfying their information requirements. As can be expected, there is greater interest in on-line systems including on-line data entry, and an accelerated effort is being made to shortcut the old approach of batch processing first followed by on-line and inquiry capabilities.

The exchange of programs and transfer of systems has proven to be more feasible and more successful than one would expect. Substantial potential savings can be effected by importing programs even when extensive re-programming is required. This was especially true with the Stanford University OASIS system which has been transported to ten other locations. The difficulties experienced in earlier attempts to transport OASIS have been resolved and, in some instances, the system has been installed and up and running in a matter of days.

In New York State, SUNY recently was offered a very sophisticated library circulation control system developed by Ohio State University and with minor revisions implemented it at the University in Albany. By so doing it is estimated that SUNY saved four to five hundred thousand dollars. Furthermore, thanks to O.S.U., SUNY obtained one of the best library circulation programs in operation.

Decision making and organizational problems add to the difficulties in the development of MIS. One sometimes cannot identify the decision maker, especially when there are many levels of users and administrators. George Weathersby structured the organizational picture by citing four myths involved in planning management information systems. First, it is usually assumed that data is neutral, but true only to the extent that many different people can use it many different ways. Second, it is widely believed that compatible sets of data are collected by institutions. The National Commission on the Financing of Post-secondary Education has encountered great difficulty in attempting to design, identify, and develop common sets of data. A third myth maintains that data collected by colleges and universities is accurate. This is not true for higher education at large. The National Center for Educational Statistics is very sensitive to certification of data or the accuracy of data. Information that hasn't been extensively tested, checked, reviewed and recertified is no longer accepted. A fourth myth concerns samples of data used for national statistics. Sampling procedures currently being used often yield extremely small data sets, and where this is the case, the use of the data is almost immoral.

In the development and use of MIS in colleges and universities one of the common problems is distinguishing between authority and responsibility. Computer people designing a system and planners and administrators defining the management information requirements often work independently. It is therefore important to identify responsibility and the authority for the total program by focusing the entire effort under one group or one department.

Communications are always a problem in discussing and trying to define users' needs and requirements. In the design of a university management information system, as many as 200 or 300 people may become involved. In New York State this has been true. Although nearly every administrator is involved from the local campus to the state systems on up to the central office, one can still question whether there has been enough faculty involvement in the design.

At SUNY, one dozen or more task forces have worked on the statewide MIS (student systems, personnel, financial, facilities, student accounting, standards, a general advisory committee, etc.). It is very difficult to keep task forces motivated. They have a sincere and a deep interest in developing uniform systems but, when it takes two years and three years instead of the one year that had been anticipated, interest wanes.

If SUNY had to do it over again, having listened to Mike Roberts and the group on the transportability of programs, more consideration would be given to the use of anything and everything that could be obtained and revised. In our case a student information system in operation at a university center was used as the base instead of completely redesigning a SUNY statewide student system. The university center student system was redesigned for the four-year and two-year college use eliminating at least 40% of systems development work. Forty percent of twelve to fourteen man years resulted in substantial savings. Universities and colleges now starting from scratch should at least look around, and find out where programs are available and in use. There is always the pride of authorship but one can design a unique MIS using many pieces or parts of existing systems without trying to do the total job.

The most time consuming but important step in the design or redesign of an MIS is the definitions of the data elements that are required. A positive aid should be the use of a very detailed checklist to make sure all the appropriate bases are covered. One instance, that occurred at Stanford illustrates the need for such a list. According to Cheryl Traver, the system staff at Stanford had nearly completed one system of OASIS when it was decided that no one would be willing to use social security numbers as identifiers. Obviously the incorporation of a multiplicity of personal codes into a half-finished system can create an interesting problem.

One of the most serious problems, even when committed to WICHE Standards, is the usual lack of uniformity and standardization by codes. For that reason the responsibility in this area should be given to a full time standards committee.

Development time and costs for MIS development are still being underestimated. Original estimates of four years of development have now been revised to eight years. Estimates of ten man years development time have often become the square of that time. An original estimate of ten man years to develop the basic system was revised, two years later, at the halfway point, to be at least 40 man years for the administrative effort in New York State. As time increases, motivation becomes an increasingly important factor. Keeping a group of people motivated over a period of years in developing total systems is difficult.

Despite problems of development time and financing, there is extreme optimism, especially about transportability of MIS programs and based on the reports presented at this conference this optimism appears to be justified.

Chapter 10

MIS Case Studies

The Effects of Management Information Systems in Colleges & Universities: A Perspective from the Statewide Level

by W. K. Boutwell
State University System of Florida

Most institutions of higher education in this country have been involved in the creation of Management Information Systems for almost a decade now, with the most serious efforts coming within the last five years. It is probably safe to say, however, that most systems are less than half complete at this point in time and several man years of effort must still be expended before such systems will be complete; if, indeed, they will ever be complete. It is time, however, that the administrators in higher education took a serious look at the effectiveness of the large amount of expenditures that have already been and will continue to be utilized to construct automated data information systems. I predict that few, if any, administrators will decide to abandon their efforts to create good Management Information Systems; going back to the old way of doing things is simply unthinkable. "You can't keep them on the farm once they have been to Paris." Nevertheless, it is good to stand off at a distance once in a while and take another look at the system one is creating.

One's evaluation of the effect of a Management Information System is directly dependent upon the perspective from which the system is viewed. You will agree, I am sure, that faculty members, deans, university presidents, chancellors of state university systems, and legislators all have a different view of the importance of Management Information Systems. Let's face it: Information is power. Normally the one who possesses the information is in a much superior position for influencing resource allocation and other policy decisions. For years, universities were in a powerful position because they possessed most of the information and state boards and legislatures were left guessing. Thus, it is not surprising to find different views about Management Information Systems and the wisdom of giving data to the "unlearned" at state and federal levels.

In this paper the effectiveness of Management Information Systems is addressed from a statewide standpoint. As the reader through these evaluations, however, he may want to think of how these evaluations might differ if they were being addressed by someone with a different perspective.

To objectively evaluate a Management Information System, one must understand the environment which necessitated the creation of the system in the first place. Thus, it is appropriate to briefly review the history of higher education in this country during the last ten years. The decade of the 60's was characterized by: 1) rapidly increasing enrollments, which placed extreme pressures on limited resources available for higher education; 2) the creation of many new institutions of higher education as many states adopted the objective of carrying education to the people; 3) general dissatisfaction on the part of the general public with the quality of outputs of many institutions (e.g., lack of relevance); and a demand for accountability in the use of public funds in all areas including education, which, in most states, accounts for about 50% of state expenditures.

Florida typifies what was occurring in most states during the 60's. In the state's universities alone, enrollments grew from 27,000 to 88,000 and the number of universities from four to nine. In addition, the community college enrollments increased from 17,000 to 122,000 and the number of community colleges grew from nineteen to twenty-eight. It was during this period of rapid and almost uncontrolled growth that the Florida Legislature realized that some overall control and planning had to be brought to higher education in the state in order to avoid vast duplications of programs and facilities. The Legislature was simply unable to cope with the competing demands of vast numbers of community colleges and universities. Thus, in 1965 the old coordinating board for the universities was replaced by a governing board and the office of chancellor as chief executive officer of a central staff of the governing board was created. The community colleges were placed in a separate division of the Department of Education. The new governing board was given total responsibility for all of the state's public universities. Thus, the chancellor, as the executive officer of the board, became, in effect, the chief executive officer of a unitary higher education system in Florida. Needless to say, the years since 1965 have been exciting ones. Universities do not voluntarily submit themselves to a system and, in most cases, do not hesitate to speak out, both publicly and privately, when the decisions of a central staff are judged to be in error. Thus came the big question, how does a chancellor's office go about the job of creating an effective and efficient state university system out of previously independent universities, each with its own political power base. Obviously, a good, reliable source of information was a prerequisite. The problem, of course, was further complicated by the fact that it was impossible for the chancellor's office to be located on all campuses. Thus, the knowledge and familiarity of programs, problems, and opportunities that come about because one is located on-campus had to be foregone. Almost all information utilized in decision making at the statewide level had to be gained through some type of a nonpersonal communication system between the local campuses and the central state offices. This is the context in which Management Information Systems must be evaluated at the statewide level.

Uncontrolled growth of both enrollments and institutions of higher education coupled with an increasing demand for accountability in the use of public funds created the need for more central planning and administrative control for higher education. Obviously, central planning and governance of institutions of higher education could not take place without the benefit of a constant flow of information concerning the attributes of students, faculty, academic programs, financing, and so on. Management Information Systems, thus, were created to supply the needed information at the statewide level. The forces leading to more centralized planning and governance came before the creation of central Management Information Systems. Some colleagues in higher education around the country argue that Management Information Systems are creating centralized decision making. However, in Florida, centralization, or at least the forces leading to centralization, came first and Management Information Systems were created to make decision making at the central level more effective.

MIS IN FLORIDA

During the latter 1960's, the Management Information System in the State University System of Florida consisted primarily of written reports, both routine and ad hoc, supplied by the universities. There was very little standardization of data items and the system, as one might guess, was a very rigid one. It was almost impossible to get comparable data from the different universities to meet the demands of the central system office as well as the demands of the governing board and the Legislature. Almost all requests for data were met with responses like "this would take two years to derive", or "we do not collect data in that form". In an effort to anticipate as many data needs as possible, the chancellor's staff gradually, over a period of four years, drastically expanded the number of written reports to be submitted by the universities until, in 1972, we were getting over 2,000 different written reports per year. There was a written report on almost every conceivable topic. Unfortunately, however, very few of these reports ever supplied the data in the exact format in which it was needed for a particular policy decision. The chancellor, the governing board, the governor's office, and the Legislature all became hopelessly frustrated with the inability of the universities and the chancellor's staff to supply the types of data and supporting analyses needed to perform their decision making responsibilites. During this time period many people began to question whether an effective State University System would ever be created and whether some other form of organization was needed to accomplish the state's objective of a well coordinated and governed State University System. It was in this context that the chancellor, in 1969, directed that an automated Management Information System be created for the State University System of Florida. In 1973, after four years of development the information system is approximately 60% complete.

Very briefly, the Florida State University System Management Information System is composed of two parts. The first part involves the creation of common data gathering and recording systems at all universities. Named UNIFTRAN (Uniform Transaction Systems) this part of the MIS has been designed primarily by the users of management information. The second part of the system involves extraction of subsets of information from each university's Management Information System to create a single statewide Management Information System. Both parts of the system have been most carefully designed so that data in any file can be matched with corresponding data in any other file. In addition, all users of management information have access to a single data bank thereby successfully avoiding the problem of having different organizations create individual Management Information Systems.

Currently at the state level, the following data files are received from each university:

- Student Data Course File
- Student Admissions File
- Instructional Activity File
- Space File
- Authorized Position File
- Faculty Activity File

These files provide the Chancellor's office with 60-65% of the information needed at the state level. There is still a void in the area of financial data where the chancellor's office continues to rely on written reports. Hopefully, however, the written reports will soon be replaced with data tapes.

EFFECTS OF MIS

What have been the effects of this Management Information System?

First the number of written reports submitted to the chancellor's office by the universities has been drastically reduced. This, of course, has allowed the universities to significantly reduce the man years dedicated to writing the reports. These resources have been reallocated to other functions within the universities and, in many cases, have been dedicated to improving planning and analysis activities in support of university administrative decisions. The second effect, and probably the most important one, is that the credibility of the State University System, including the individual universities, before the Legislature has been greatly strengthened. For the first time, the chancellor's office has the ability to respond to most of the Legislative requests for information needed to analyze alternative legislative policies regarding higher education. Three years ago the central system office could not: produce a class size analysis for all of the public universities, furnish information concerning the number of tenured and non-tenured faculty, produce space utilization data, nor provide information concerning the rank of the faculty members teaching freshman courses. Examples are plentiful. Today the chancellor's staff can respond to almost all data requests with only a couple of weeks' notice. The Legislature is now able to reach many policy decisions based upon full sets of data and, more importantly, Legislators have gained confidence in the decision making ability of both the chancellor and the university presidents. A third major impact of the Management Information System has been a significant improvement in the availability of data for decision making within the universities. It is not uncommon for vice presidents and deans to rely upon computer printouts of faculty productivities, average class sizes, projected enrollments based upon induced course load matrices, and costs per student credit hour as a base for determining the allocation of resources among colleges. Nor is it uncommon to find deans and department chairmen doing the same thing.

These indicators should not be misleading. There are still problems: data voids in the data banks; coding errors; and keypunch errors. Often data in the Management Information System is still not available in the form in which it is requested or needed by administrators. However, life is far less frustrating these days than it was two and three years ago. Resource allocation and other policy decisions can now be recommended to the chancellor and the Board of Regents based upon objective analyses of detailed data. Most staff time now is spent in analyzing data for the purpose of reaching policy decisions, instead of hunting some small set of data somewhere that might possibly be used to support a decision that has already been made.

Management Information Systems have drastically improved the level of decision making at the statewide level in Florida during the past four or five years. This improvement should continue as staff refine and further automate the Management Information System. Similar improvements have already occurred at the university level.

MIS: The Ugly Duckling

by John W. Gwynn
University of Wyoming

HISTORICAL SETTING

The concept of MIS came a few years ago when management realized that data being used for daily operations could substantially aid in decision-making, planning, and analysis. However, it soon became apparent that needed information known to be in machine readable files was not readily available and that which was available usually came too late to be of any use. Administrators found that existing systems could not respond to an ever-changing variety of inquiries. To obtain information it was necessary to communicate with a programmer and then wait for a program to be coded, debugged, and submitted for production processing. Besides being late, the information finally received was low in quality and utility. Little could be done to improve the situation because existing application systems were difficult to modify and required high maintenance just to keep them in operation. Data needs in the management environment were found to change faster than traditional programming methods could accommodate.

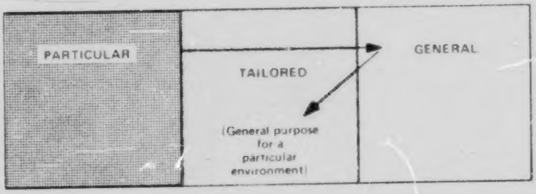
With high anticipation the first generation management information systems were born. The attack on the problems was very direct as systems were designed to put the manager in direct contact with his data base. One of the first problems encountered during the early development was that existing files were inconsistent and contradictory. Similarly titled data elements in differing systems were found to have conflicting definitions and formats. Soon, files were integrated, put on-line and time-shared systems began to respond to inquiries.

Although possessing many useful and desirable features these systems introduced about as many problems as they solved. Those with sufficient sophistication to do the job required large, powerful, and expensive hardware. Though addressed to solve a new breed of problems, they were programmed using old techniques and methods. As a result they required a large maintenance staff, they were slow to repair, and modifications were time-consuming and costly. The software was so complex that error-producing conditions could not be recreated to allow repair. First-generation systems were found to consume vast resources and were expensive to develop and install. Because they required a large volume of on-line data storage and consumed significant amounts of computer time, production batch work was seriously degraded and there was no way to reschedule peak loading. File integration and data base development has brought out new dimensions to concerns for file security and data integrity. As the visibility of errors and malfunctions has broadened to include a community of non-technical on-line users, personnel problems in the computer center have grown.

Perhaps the final blow to the first attempts to build MIS systems was the fact that they were not transferable and could not be easily moved to the newest and latest hardware. Installed late, over budget, below expected performance, exhibiting low computer utilization, difficult to expand, and tightly married to particular hardware, these systems have become software's ugly duckling.

Although the summary of events discussed in the preceeding paragraphs describes experience with management information systems, it closely parallels early computer hardware development and the development of compilers and applications programs. History has seen both hardware and software move from the very particular to the over-generalized and finally to the "tailored" (See Figure 10.1). If MIS systems development follows this pattern it should be valid to conclude that one can learn enough from the first and second generations of MIS software to develop a greatly improved and viable third generation.

Figure 10.1 Computer Software Development



PLANS AT WYOMING

The University of Wyoming is committed to the development of a management information system which is currently being designed. How well the new system is able to solve the problems mentioned above remains to be seen. However, much discussion has preceeded the direction of current thought and concepts and methodology are being specifically applied toward particular goals. The new system is going to be a synthesis of many new developments in an attempt to apply the results of experience and research in a practical way.

Encouraged by the recent explosive success of turnkey data processing applications on minicomputers [1], the University of Wyoming will be using minicomputer hardware. In addition, the pioneering work at University of California-Irvine by Professor Farber and his associates confirms the viability of a network of small computers [2,3,4,5,6]. The Wyoming network will employ off-the-shelf hardware which is modularly expandable from a single minicomputer to many. Such low cost, high power hardware composed of independently connected processing nodes offers high reliability, true modularity, modest programming requirements, low overhead, and improved response.

Features to be included are derived from Stanford University's OASIS[7], Dartmouth College's Project FIND[8], MRI Systems, Inc. System 2000[9], Cincom Systems, Inc. TOTAL[10], IBM's IMS[11], and Software Ag's ADABAS[12].

The software will be written in an intermediate language[13] as was done at Dartmouth on Project FIND. In this manner it becomes essentially hardware independent and makes it possible to rapidly take advantage of the newest, fastest, and cheapest technological developments as they occur. By programming well designed modules the software features become building blocks which may be assembled into a variety of actual systems. Thus, differing needs may be satisfied by selecting only the features required. This building-block approach will make it possible to begin with a very small hardware configuration and expand as the situation demands. For example, an initial installation might be a single minicomputer system with software offering generalized inquiry, report generation, and file maintenance[14]. The next step could be the adoption of software to support tailored applications where the generalized services are available part of each day and the tailored applications are available part of the time.

A third increment could be to add more hardware which would allow some network nodes to be dedicated to generalized services and some to be dedicated to tailored applications. The final addition would be to include full software support to permit completely generalized resource allocation for distributed computing and multi-processing.

Experience has shown that maintenance of traditional application systems consumes a significant portion of staff time and resources. Most generally it will be found that these high-maintenance systems were designed to be installed and left alone. Yet the actual environment is one of constant change. Although inflexible programs may be modified, the cost is high and the results are often unsatisfactory. New systems design and programming techniques utilizing both hardware and software modularization concepts are urgently needed. Systems design to anticipate change are part of the new breed of MIS. The University of Wyoming will employ as many of these methods as are presently known to work. All software from the basic operating system to the applications will be modular[15,16,17,18] and programming techniques will be used which allow on-line interactive programs to be run in batch using the same program code. By design the systems will use a language which makes it possible to obtain new hardware without disrupting operations. Files will be disengaged from the applications software making it easier to reorganize the data or change the structure and peak loading will be distributed among independently operating nodes of the network.

Large, centralized data processing facilities suffer from the fact that when the system goes down everybody and everything is stopped. Because delayed production backs up into later work, it is often some time before normal operations are resumed. A distributed system, on the other hand, makes it virtually impossible to completely stop operations[19]. In a fully distributed system component failure may degrade service but rarely stops it entirely. A partially distributed system may be designed such that component failure stops a single user or a single service to all users. The first version of the Wyoming MIS will be designed to insure if one node of the network fails only a single user will be down. In other words, control and files will be distributed whereas the traffic on the network will be limited to data retrieval. This simplification reduces both hardware and software complexity for the "phase one" implementation.

This system will probably be of value to other institutions. The emphasis on hardware independence and turnkey generality will make portability an additional benefit.

COST BENEFITS

Aside from the technical details of the hardware and software, one of the main considerations when pondering an MIS is the cost. Although much of the advantage of a good MIS is in quality improvement, not enough attention has been given to savings. Although difficult to measure, some very real savings may be realized. Included in the list are:

- Lower total hardware cost
- Lower cost hardware maintenance

- Lower cost development and maintenance of software
- Displacement of multiple programs with a single system
- Lower cost modifications
- Less critical and less damaging personnel turnover
- More relevant data
- Peak-load distribution
- Shorter application program lead-time
- Increased staff productivity

Because software is simpler, one realizes lower cost for software development and maintenance. At the University of Wyoming the new MIS will directly displace over 400 batch inquiry and report generation programs. With that displacement will go the attendant resource expense for maintenance, modification, documentation and scheduling. Personnel turnover is less critical because a well-modularized system is more easily maintained by someone other than the original author. Much of the data which is presented on report from traditional systems is not used or must be re-worked by hand to be useable. An on-line interactive MIS permits direct access to exactly the information required and eliminates the expense of producing almost correct or never used data. Although peak load distribution is a problem with first generation MIS systems it need not be in a new system. Emergency or little-used over-capacity hardware is not needed. Finally, shorter application program lead-time can be achieved. An MIS facilitates new applications work by providing many of the needed services directly.

REQUIREMENTS FOR A SUCCESSFUL MIS

There are ten requirements which a successful MIS should include. This list will serve as a guide during the design of the University of Wyoming MIS and could be a set of criteria against which any system purported to be an MIS could be compared or measured.

1. Easy to implement and use
    - Automatic take over of all existing files and data and integration into the data base
    - Transparent data structure, ordering, and format
    - Various hardware and terminal types
2. Easy to change and designed to change
    - Modular software
    - Transferable: written in an intermediate language
    - Transportable: able to move system to a variety of locations
    - Data field expansion or cancellation at any time
    - Data disengagement
    - Data base growth without reorganization
    - Incremental hardware expansion

- Incremental software feature expansion

3. Broad performance range
   - Automatic file information coupling (indirection)
   - Simultaneous on-line and batch
   - Automatic storage device administration
   - "Forgiving" (probability retrieval, phonetic searching, etc.)
4. High system availability
   - Modular hardware: not everybody down at once, easy maintenance
   - Multi-tasking, time-sharing
   - Reliable hardware and software
5. Maximum system and data integrity
   - Efficient backup/restart
   - Multilevel data security
6. Optimum performance and efficiency
   - Retrieval speed independent of data volume (random access)
   - Sequential processing speed comparable to stand-alone applications
   - Search strategy optimized from data base semantic content
7. Complete information system support
   - Generalized inquiry
   - Generalized report generation
   - Generalized update
   - Tailored applications support
   - Application system debugging aids
   - Utilities
8. Low Cost
   - Hardware acquisition and maintenance
   - Software acquisition and maintenance
9. Responsive and Adaptive
   - Responds to unforeseen requests
   - Easy to modify or add new features
   - Increases staff productivity
10. Low Maintenance as a Turnkey system

By employing new techniques, new technology, and new design principles it should be possible to implement a management information system which has all the features and power of former systems but with fewer problems and about one order of magnitude reduction in cost.

REFERENCES

1. There are many articles reporting successful turnkey installation of minicomputers with applications software. One excellent selective bibliography containing references to such articles and a list with addresses of eight turnkey minicomputer system suppliers is available on request to subscribers of *EDP Analyzer*, 925 Anza Avenue, Vista, California 92083.

2. "Where Minicomputers Do All of the Job". *Business Week*, July 14, 1973 p. 38K-38L.

3. Farber, D.J. "A Survey or Computer Networks". *Datamation* 18, 4, April 1972, p. 36-39.

4. Farber D.J. and F. R. Heinrich. "The Structure of a Distributed Computer System – The Distributed File System". *Proc. International Conference on Computer Communications*, October 1972, p. 364-370.

5. Farber, D.J., M. D. Hopwood, and L. A. Rowe, "Fail-Soft Behavior of the Distributed Computer System". Technical Report #24, (Department of Information and Computer Science, University of California, Irvine, California) November 1972.

6. Farber, D.J. and K. Larson. "The Structure of a Distributed Computer System – The Communications System". *Proc. Symposium on Computer-Communications Networks and Teletraffic*, (Microwave Research Institute of Polytechnic Institute of Brooklyn) April 1972.

7. "OASIS – General Introduction". (Project INFO, Stanford University, Palo Alto, California) 1973.

8. "Project FIND Manual". (Project FIND, Dartmouth College, Hanover, New Hampshire 03755) 1973.

9. "System 2000 – General Information Manual". (MRI Systems Corporation, 12575 Research Blvd., Austin, Texas 78766) 1972.

10. Information about TOTAL and ENVIRON/I may be obtained from CINCOM Systems Inc., The Registry, Suite One, 6350 LBJ Freeway, Dallas, Texas 75240.

11. Check with your local IBM sales office.

12. Schnell, Peter M. "ADABAS Adaptable Data Base System – Introduction", (Software AG North America, 12124 Bassett Lane, Reston, Virginia 22070) 1972.

13. Project FIND uses an extended Dartmouth BASIC as the implementation language. Other projects have used standard high-level languages such as PL/I and FORTRAN or specially developed intermediate languages such as L. LOWL, PL/36, and MOLSUF. A brief report of experience in developing portable software is "Levels of Language for Portable Software", *Communications of the ACM*, Vol. 15, Number 12, December 1972, p. 1059-1062. The DCS project at UC Irvine uses a language called MOLSUF (see references 2, 3, 4, 5, 6 above).

14. See references 7, 8, 9, 10 and 12 for illustrations of single minicomputer systems with software offering generalized inquiry, report generation, and file maintenance.

15. Armstrong, Russell M. *Modular Programming in COBOL*. (John Wiley and Sons, New York) 1973.

16. "Modular COBOL Programming", *EDP Analyzer,* Vol. 10 Number 7, July 1972.

17. Parnas, D. L. "On the Criteria to be Used in Decomposing Systems into Modules", *Communications of the ACM,* Vol. 15, Number 12, December 1972, p. 1053-1058. "On the Criteria to be Used in Decomposing Systems into Modules", Technical Report CMU-CS-71-101 (Carnegie Mellon University Dept. Computer Science, Pittsburgh, Pa.) August 1971. Other reports on this topic from the Department of Computer Science at Carnegie-Mellon are: "Some Conclusions from an Experiment in Software Engineering Techniques"; "Information Distribution Aspects of Design Methodology"; "Use of the Concept of Transparency in the Design of Hierarchically Structured Systems"; "Response to Detected Errors in Well-Structured Programs"; "A Program Holder Module"; "Design and Implementation of a Multi-Level System Using Software Modules"; "Implications of a Virtual Memory Mechanism for Implementing Protection in a Family of Operating Systems".

18. Dijkstra, Edsger W. "Notes on Structured Programming," T. H. – Report 70-WSK-03, (Department of Mathematics, Technological University Eindhoven, The Netherlands) April 1970.

19. Farber, D. J., M. D. Hopwood, and L. A. Rowe. "Fail-Soft Behavior of the Distributed Computer System". Technical Report #24, (Department of Information and Computer Science, University of California, Irvine, California) November 1972.

Project FIND*

by John S. McGeachie
Donald L. Kreider
Dartmouth College

Project FIND (Forecasting Institutional Needs at Dartmouth) has been established at Dartmouth to make institutional data readily accessible to administrative officers and faculty members through the facilities of the Dartmouth Time-Sharing System (DTSS). A concomitant goal is to develop models of the operation of the institution to facilitate long-range planning by providing quantitative estimates of the effects of policy changes.

Thus Project FIND has a three-fold purpose: 1) to provide a tool, requiring minimal computer expertise, through which administrators and faculty members can obtain current information on the college's finances, students, staff, and physical facilities; 2) to provide a universal format and language through which members of the college community can manage their own private data; and 3) to provide a capability for making projections of expense revenue trends, tenure ratios, student choices regarding major fields of study, requirements for academic space and student housing, and so forth.

By October 1973 the first and second objectives had been met. An easy-to-use interactive information retrieval and analysis system, also called FIND, is available through terminals connected to the Dartmouth Time-Sharing System. The FIND system has been operational for a year and is used by personnel in the offices of the President, the Budget Officer, Personnel Administration, Alumni Affairs, and Student Affairs.

---

*Much of this paper has been taken from the Project FIND Interim Report, by D. L. Kreider and J. S. McGeachie, October 1973.

Significant progress has been made toward the third goal, a capability for developing institutional models, and the FIND system can now also operate as a program-driven data retrieval and analysis system, controllable through programs written by the user in the BASIC language. Programs can be written which use the full power of the FIND system as well as the full computing facilities of the Dartmouth Time-Sharing System.

The purpose of this report is to discuss some of these applications of the FIND system and to indicate directions that further development will be taking.

THE FIND SYSTEM

Project FIND was officially launched in March 1972, when Dartmouth College President John Kemeny gathered a group of college officers and consultants at Dartmouth's Minary Conference Center to discuss the purposes, a timetable, and a tentative design for an information retrieval and modeling system.

The first version of the information retreival system was completed by June 1972, with the bulk of the programming having been done by a group of undergraduate students in a course taught by President Kemeny. Two of these students continued to work on the project during the summer as members of the FIND staff, and by September 1972, the original programs had been refined, new programs had been added to allow simple statistical analyses and cross-tabulations, several data bases covering faculty and administrative officers had been brought into existence in a format specifically designed for Project FIND, and the first modifications in the FIND system that would make it accessible to users' programs (for modeling purposes) were completed.

One of the fortuitous aspects of the modeling capability was that users across the campus were able to write programs in BASIC to manipulate their own data available through FIND. Some of these user programs were subsequently studied by the Project FIND staff and incorporated into the official FIND system for everyone to use. Thus any member of the Dartmouth community is a potential contributor to the FIND system, and the evolution of the system in directions needed by its users is ensured.

A great deal of effort was devoted to making the FIND system easy to use, and a beginning user needs to know only a few commands in order to access FIND. The most important of these commands is RETRIEVE, which specifies the general area of the user's inquiry. For example, the command "RETRIEVE FACULTY, SURNAME, DEPT, RANK, SEX," indicates that the user is interested in the faculty data base and wishes to examine the surname, departmental affiliation, rank and sex of its members. These four qualifiers are called attributes; a typical data base may have over one hundred attributes for almost a thousand members. The attributes associated with the RETRIEVE command are collectively known as the working data base. A user's working data base is usually very much smaller than the full data base.

If the user wishes to focus his inquiry on a subset of the faculty, he may use the SELECT command, which restricts the system's attention to members who meet the user's specified criteria. For example, the command, "SELECT DEPT = "HIST"," would narrow the scope of the inquiry to include only members of the History department.

The SORT command may be used to reorder the working data base elements in either ascending or descending alphabetical sequence based on specified attributes. For example, the command, "SORT RANK," will sort the previously retrieved data elements in ascending order by academic rank. To obtain a descending sequence, one merely precedes the attribute name by a hyphen, as in, "SORT -RANK. -SURNAME," which sorts the working data base in reverse alphabetical sequence within descending order of rank.

The PRINT command is used to display the working data base after it has been reduced and rearranged by SELECT, SORT and other commands. The example in Figure 10.2 shows how to obtain a list of Admissions Office administrative personnel, in descending order by appointment date. By using a FORMAT command one obtains a printout which fits on that page.

Statistical functions are available through commands such as XTAB which performs cross-tabulations. (See Figure 10.3) For example, to obtain a breakdown of rank versus sex for the previously selected members of the faculty, the command, "XTAB RANK, PTS, SEX, STD, %" where "PTS" means starting points and "STD" means standard groupings, would enable the user to specify up to eight ranges for the attribute rank, after which FIND would display percentage figures for each sex within the specified ranges. Appendix A contains a list of the available FIND commands, and Appendix B lists data bases available in December 1973.

INSTITUTIONALIZATION OF FIND

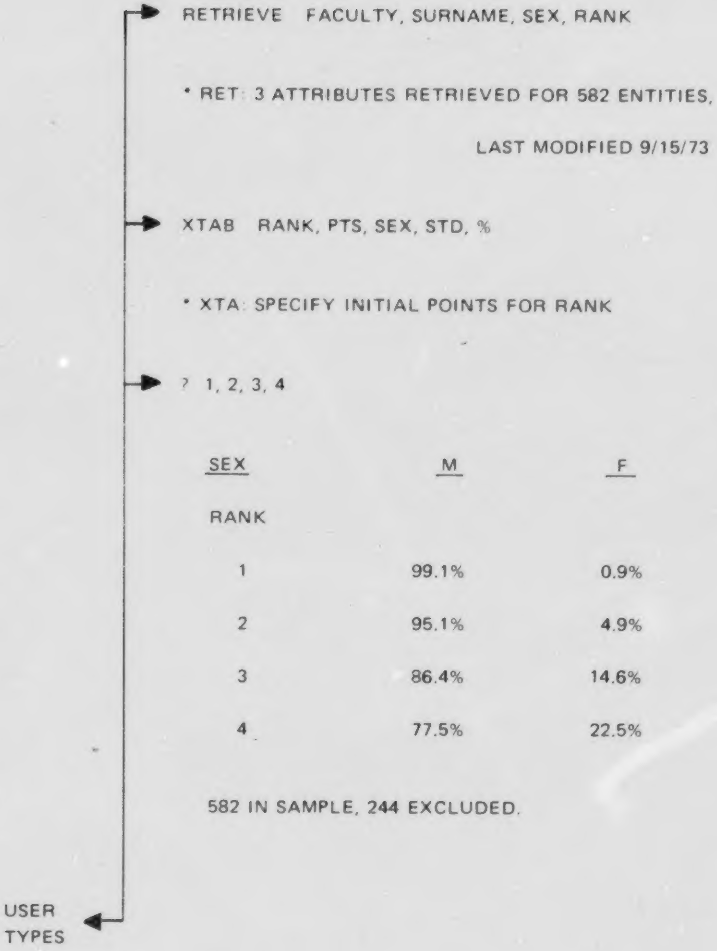
The first data bases that were accessible to the FIND system were constructed on an ad hoc basis for experimentation. Thus, although they served this purpose well, they could not be depended upon for up-to-date accuracy. They were not yet the "official" data bases of the college.

Throughout the academic year 1972-73 a major effort was directed to four problems that required solution before the institution could "turn

Figure 10.2 FIND Lists Personnel by Appointment Date

```
DARTMOUTH TIME-SHARING
LINE 231, ON AT 12:34 12 DEC 73, 119 USERS
USER NUMBER -- 123456, MMMMMMMM
NEW OR OLD--OLD DPCLIB***:FIND
READY

RUN

FIND         (COMPILED) 12 DEC 73  12:38

FIND HERE!

? RET ADMIN, SURNAME, INITIALS, GRADE, SEX, DEPT, HIREDATE
*RET:  6 ATTRIBUTES RETRIEVED FOR 271 ENTITIES
       LAST MODIFIED 12/07/73
DONE

? SELECT DEPT = ADM
*SEL:  14 ENTITIES SELECTED
DONE

? SORT - HIREDATE
DONE

? FORMAT  W51
DONE

? PRINT INITIALS, GRADE, SEX, DEPT, HIREDATE

INITIALS    GRADE    SEX    DEPT       HIREDATE

   P          1       M     ADM         730906
   J          1       F     ADM         730901
   G          1       F     ADM         730821
   W          1       M     ADM         730814
   J          1       M     ADM         730430
   P          1       F     ADM         720901
   L          4       F     ADM         720201
   L          2       F     ADM         710819
   W          5       M     ADM         690701
   T          3       M     ADM         680901
   G          5       M     ADM         640935
   T          6       M     ADM         631001
   J          4       M     ADM         620901
   T          7       M     ADM         360701
DONE
```

Figure 10.3 FIND Generates Frequency Tables



```
RETRIEVE   FACULTY, SURNAME, SEX, RANK

* RET: 3 ATTRIBUTES RETRIEVED FOR 582 ENTITIES,

                         LAST MODIFIED 9/15/73

XTAB   RANK, PTS, SEX, STD, %

* XTA: SPECIFY INITIAL POINTS FOR RANK

? 1, 2, 3, 4

  SEX              M            F

  RANK

    1            99.1%         0.9%

    2            95.1%         4.9%

    3            86.4%        14.6%

    4            77.5%        22.5%

  582 IN SAMPLE, 244 EXCLUDED.
```

over" its data keeping functions to Project FIND.

Updating of files. An effective and simple capability for creating new data files and for keeping old ones up-to-date had to be developed. Ultimately, the accuracy of FIND's data could be guaranteed only if the many officers of the college responsible for *generating* such data could also be assigned the responsibility for *updating* the FIND files. Project FIND's data bases had to become the primary and "official" repository of the institution's data, not merely copies of it.

By the summer of 1973 the required capability for updating files was in existence. From any computer terminal of the Dartmouth Time-Sharing System, a user can enter the FIND system and issue the command "UPDATE" He can then effect a succession of changes in any data files under his control by merely typing the additions, deletions, and modifications directly to the computer. Alternatively he or she can accumulate a log within the computer of changes to be made to an official institutional data base, and can have such changes effected automatically by making an authorizing call to the Data Processing Center.

Given the existence of the UPDATE system, Dartmouth assigned official responsibility and accountability to the Office of Personnel Administration for the maintenance of all FIND data bases relating to employees of the college. As of October 1, 1973, the FIND personnel files are the official files of the college and are kept up-to-date on a day-to-day basis. The FIND files supercede all of the previous ad hoc files that were often inaccurate.

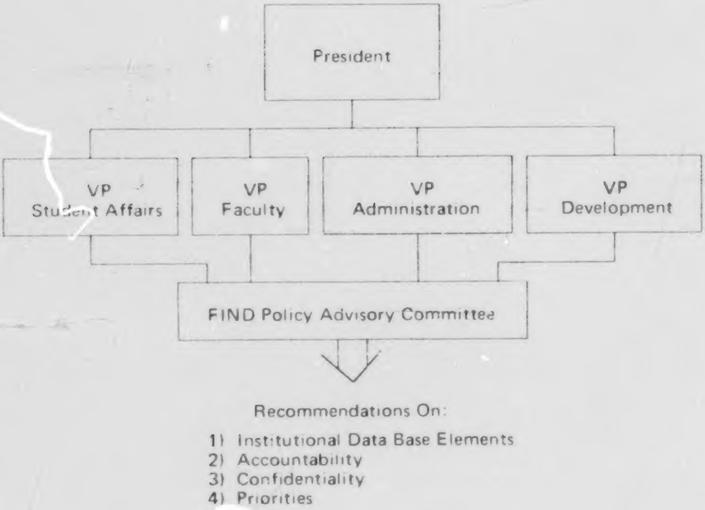
Selection of data elements. It was necessary to make a major study of what data is needed by each office or part of the college to determine who generates and uses such data, and to learn how data flows from one part of the college to another. The Office of Institutional Research and Analysis made such studies in the areas of students and personnel which produced an inventory of all data needs, identified the redundancies in the collection of data, and pointed out inconsistencies in the definition and use of data by different offices. The recently established Personnel Data Base was made possible as a result of these studies. And the Student Data Base, currently under construction, relies heavily on the inventory study.

Protection of information (confidentiality). It was necessary to develop the technical capability for protecting confidential and sensitive data from unauthorized and unintended use. By the late fall of 1972, a highly ingenious security system was implemented, using a technique of encyphering sensitive data. The user effects the encyphering through a password that he himself gives and which is retained nowhere in the computer system. So effective is this technique, that even a minute difference in the password given will result in totally unrecognizable gibberish. If a user loses his own password, not even the FIND staff can help him out. The only possibility would be to recreate the data base. Thus, the technical aspects of security have been solved.

Data Base Accessibility. It has become necessary to address the question of who is entitled to access data in the FIND files. This is, of course, a matter of college policy, and it is being resolved through extensive consultation with college officers responsible for collecting data. The final policies adopted must accommodate the rights of individuals who voluntarily provide information to the institution, yet the purposes of Project FIND as a planning tool must not be unduly restricted.

An institution-wide advisory committee for Project FIND has worked on such matters throughout the 1972-73 year under the chairmanship of Vice President Donald Kreider. This group approved in principle last May a set of recommendations from the Office of Institutional Research and Analysis on which data should be considered institutional data (as opposed to private data relevant only to one office). The committee also approved in principle a set of guidelines on confidentiality, accessibility, and accountability. Documents on all such matters considered by the FIND staff and the several advisory committees are kept on file and should be useful to other institutions working their way through the maze of technical, political, and organizational problems entailed in the development of management information systems. This aspect of the Project FIND organization is shown in Figure 10.4.

Project FIND Organizational Structure Figure 10.4


President
VP Student Affairs
VP Faculty
VP Administration
VP Development
FIND Policy Advisory Committee
Recommendations On:
1) Institutional Data Base Elements
2) Accountability
3) Confidentiality
4) Priorities

REPRESENTATIVE DATA BASES

The Budget Data Base. The budget of the institution will constitute a major data base under FIND that will tie together all the other data bases including personnel, students, space, in the modeling process. Substantial progress has been made towards creating this data base.

Appropriate identifiers for the more than 20,000 college budget accounts have been agreed upon. Attributes that provide both current and historical information about each account have been recast into FIND format. Attributes that relate each account to the college's administrative structure and to other data bases like personnel and space have been included, and an attribute has been included that achieves a crossover from Dartmouth's organizational budget format to the functional categories proposed by NCHEMS (National Center for Higher Educational Management). Dartmouth is a participating member of the NCHEMS project.

A test budget data base has been loaded into the FIND System, and FIND system programs have been used to manipulate this data and to identify and correct unanticipated problems, such as difficulties with the method of encoding the data. Programs for producing the variety of budget reports needed by operating managers were written and debugged, and additional software requirements of the FIND system needed for efficient handling of very large data bases were spelled out and recommended to the FIND staff; these are currently being developed.

The full institutional budget has now been recast in FIND format and has been successfully loaded experimentally. Declaring it to be the "official" budget data base awaits only the required addition to the FIND system of two technical features: 1) a capability of retrieving efficiently a part of a data base as opposed to the entire data base; and 2) a catalog feature that will enable each college officer to retrieve just his part of the budget without access to the full budget. The development of these two features has the highest priority of the FIND staff in 1973-74.

The inclusion of an NCHEMS attribute in the budget data base makes it immediately possible, under the FIND system, to recast the Dartmouth budget in a form suitable for comparison with other institutions that cooperate with NCHEMS. The next crucial step in the long-range development of FIND is to produce further attributes that permit the reformulation of the budget along functional lines defined especially for Dartmouth. A major effort has begun on the development of alternative budget models which are expected to be operational by late 1974.

Several uses of Project FIND have developed spontaneously. Three examples serve to illustrate how the accessibility of the system lends itself to innovation through use by the college community. The Dartmouth College Art Gallery maintains an inventory and monitors the distribution of its approximately 14,000 art objects through the FIND system. The file of art objects is updated by the Art Gallery staff rather than by the Data Processing Center.

A complete inventory of faculty involvement in committee and student advisory capacities is maintained in a FIND data file and is updated continuously by secretarial staff in the Dean of Faculty's Office. When this information was kept by the Registrar's Office for several years utilizing computer programs, it was accessible to only one person on the campus. In September 1973, the information was translated into the FIND system and became immediately available to all faculty committees and department chairmen through individual computer terminals.

During the summer of 1973, a request from the President for historical information on college investments resulted in the creation of a FIND data base as the best vehicle for transmitting the required information to the President. Although a small data base, containing 21 different attributes for 62 different entities, the true value of the new system lay in not restricting the President's knowledge to answers to explicitly formulated questions. He is now in a position to experiment with the investments data base and to quickly answer original questions.

BUDGET AND TECHNICAL STAFF

Project FIND has been awarded $275,000 for the period March 1972, through December 1974. These funds have been provided by the Rockefeller Brothers Fund ($100,000), an anonymous alumnus gift ($100,000), a grant from the Exxon Education Foundation ($60,000) and a grant from the Mobil Foundation, Inc. ($15,000). Dartmouth College is additionally contributing $140,000 of its own resources during the period covered by the project.

The project's technical staff consists of six people: a director; three full-time specialists in the areas of software development, data base management and user services; and two student programmers. Project FIND reports to the Vice President and Dean for Student Affairs, Donald L. Kreider, who is also chairman of the FIND technical advisory and policy committee.

Evolutionary Tendencies of MIS at the University of Houston

by Ralph C. McKay
University of Houston

The term "Management Information System" or MIS is very nebulous. As it relates to computing, any system that provides information to management can qualify as a MIS. The real concern is to what degree such a system becomes a tool to accomplish management objectives.

It has been the experience at the University of Houston that a MIS is a product of evolution as opposed to revolution. Considerable planning must go into the initial implementation of the process, but a true MIS is not achieved until the base of information is generally available to management and has been accepted by management as an intrinsic part of the decision making processes. A MIS is a state of the mind, especially of the mind of management. This fact is too often overlooked in discussions on Management Information Systems.

This paper discusses the experience of the University of Houston in evolving to its present stage of implementing a MIS and in particular, how the university is affecting the transition from integrated data systems to a MIS. For purposes of this discussion the term 'system' means a given collection of programs, working with a base of information to accomplish fairly specific objectives. Examples are a payroll system, an accounting system, a student records system, and so on. The term MIS means a number of 'systems', integrated with each other for the purpose of furtherance of the end product.

UNIVERSITY OF HOUSTON MIS

The University of Houston has implemented systems to perform most of its accounting, personnel, payroll, faculty service, student aid, and similar personnel/fiscal functions. The present systems are anywhere from one to six years old. All of these applications are on-line, randomly accessed bases of information, with updating being performed simultaneously by multiple users, located in multiple offices. At the same time data may be accessed via batch programs and the timesharing applications. There are twenty-seven data files within the present fiscal/personnel/aid system plus six system files that serve to describe the data files, provide security, and so on. The vendor package used to maintain these systems includes routines for security, recovery, an interactive processor, and a report writer. The interactive processor is the primary update tool for the user offices with the exception of the accounting system, where locally-coded processors are more useful. The current size of the data base is approximately 30,000,000 characters.

The Financial Information Systems Department is part of the University Computing Center. Although technically located in the Academic Affairs organization, the department serves the entire University community. The Computing Center facility is utilized in a multi-campus environment as the primary source of computer support for the instructional programs and research efforts, as well as administrative data processing. The center operates an open shop insofar as academic support is concerned but a closed shop for most administrative data processing. Both local and remote batch service is rendered as well as timesharing.

Three different phases are apparent in the evolution to a MIS at the University of Houston: 1) system implementation; 2) system enhancements; and 3) attaining MIS status. However, none of these phases have a definite beginning and end, but tend to blend together and are cyclical in nature.

INITIAL IMPLEMENTATION

During the initial implementation of a given system, conversion is made from tape systems, card systems, manual systems, and in some instances, no defined system at all. The approaches utilized by the University in designing various systems range from University-wide committees, fiscal offices committees, one office direction, one person direction, to computing staff direction. There has been no set pattern. Approximately six years ago, the Vice President for Management Services was instrumental in the formation of a University-wide committee to act upon the information requirements of the University. This committee, composed of upper and middle management from all segments of the University, appointed sub-committees to review more specific requirements, normally based on old system lines. This process did not result in system specifications, but did prove invaluable as a point of reference for more detailed analysis and design.

The primary concern of management at this early stage was to insure timeliness and adequacy of existing batch-generated reports. The need was expressed for new products but invariably no serious consideration was given to on-line products with any degree of flexibility in the hands of the user. Although such possibilities were recognized, they were relegated to the future.

Several changes in the attitude could be recognized during this stage. First, lower management and, to a lesser degree, middle management started taking more interest in what information was available. Source documents, containing a multitude of data, much of which was never captured into the system, were often re-discovered by the users who became more aware of potentials. On the negative side, managers became frustrated because they recognized the availability of new products, but the programming staff couldn't move fast enough for them.

Some of the functional reorganizations in this phase were very healthy. Because inputting was decentralized by system approach, purchasing agents did their own encumbering, the Budget Office implemented appropriation decisions, and Personnel entered biographic data in a system. Earlier, one office had acted as the data entry point for several offices. Although data entry functions were decentralized in the new systems, little, if any actual reorganization occurred.

SYSTEM ENHANCEMENTS

The enhancement phase often coincides with the final stages of implementation. Usually system enhancements involved improvements within a given system. Eventually, however, the desire for improvement brought about a need to cross system boundaries either to improve data collection mechanisms and/or to improve the adequacy of the information.

During the enhancement phase individuals often start changing their attitudes toward the system. For the first time the system was reaching beyond its original bounds. This often means more than one organization is affected and new exposure brings difficulties. When a system is fairly self-contained, the status of the information is controllable by a single person or organization. When expansion involves more persons and organizations, new working relationships must evolve with the programming staff acting as consultants to all parties.

At the University of Houston middle management continued to be somewhat frustrated during the enhancement phase. One common difficulty was the lack of standardization of classifying items. For instance, in the new accounting system, classification of expenditures could be quite detailed if one so elected. The preceding classification scheme provided for only six divisions for control purposes. Expenditure classification for control purposes was reduced to two, personnel services and non-personnel services. Up to 2,000 categories were avialable for report purposes only. The administration elected to group these along the original six division lines with the departments having the option to choose their own refinements if they so desired. Although the departments like this, for they could classify as they desired, the administration could not perform any more analyses on their expenditures than they could earlier. Middle and upper management who wanted to perform new analyses, knew they had the mechanism for it, but couldn't get meaningful results. Certain types of stand[illegible]tion have been imposed upon the departments as a result. It now app[illegible] that more standardization will be installed in the near future.

Similar problems occurred with other classification schemes. The capability was there. For some reason, management failed to appreciate or foresee the need for more refinement. Later, when the more basic problems were reduced, they came to desire the more detailed classification.

Another type of enhancement that took place in phase two is integration of new modules into the system to enlarge the base of information. One of the prime driving forces during this phase of evolution at the University of Houston was the manager in the fiscal offices who wanted new information for his own functional area. As upper management desired more information and more refinement of such information, middle management started placing a new, more sophisticated type of requirement upon the computing personnel. This "desire-for-information explosion" normally exceeded the programming staff's ability to stay up with demand. The system was now maturing into it's final phase and becoming a MIS.

ATTAINING MIS STATUS

A database is not qualified to be called an MIS until the mechanism can be available for responsiveness to the user, perhaps to the entire University community. Stored information is of no value if it cannot be made easily available and is wanted by the user. The University of Houston is reaching the status of a true MIS in 1973.

In 1972 the University embarked on a program to place in the hands of users as fast as practical, sufficient tools to enable him to "do his own thing" with data in the management information systems. The integrity of the data has been, and will most likely continue to be, secured by controlled, predictable methods of data maintenance.

THE HOUSTON DATA MANAGEMENT PACKAGE

The primary fiscal systems of the University are maintained via the same data management package. Security by user is maintained along file lines. Read-only permission may be given or read-write capability may be extended. No mechanism is present for security at the record level. With these capabilities and limitations, various central administrative offices have already been given three important tools for securing information: interactive query processor, interpretive report writer; and read only capability via a higher level langauge.

In high frequency instances of query needs, a number of specialized programs have been developed to be executed by the user in the timesharing mode to ascertain the status of given records. These programs are directed primarily at specific inquiry such as a request to print the records of a given employee, print the status of a given account, or print a skeleton of a given account's activity since the start of the current month. These programs are utilized as an aid in the production flow and as a means of responding to specific questions as to status. Quite often, however, the question is being asked by an individual outside the fiscal office. Consequently expansion of this library approach is continuing.

In early 1973, computer center staff conducted training seminars on each of three basic information systems and on the use of a report writer. Although some middle management attended, the lower management and clerical employees comprised the majority of thirty individuals attending each series. Each series taught the use of this tool with a given system with some explanation of use across system lines. The report writer package had been developed to the point that the user did not need to be aware of all of the job control language of the system. If the user knew how to log-on and execute a given program, his or her hand was held from that point on. Furthermore, the basic package is simple enough that a non-technical person can feel comfortable with it. All of those attending, had on-going responsibilities with the system on which they were being trained.

The interpretive report writer has been a well received tool. Currently the center averages about four jobs a day through this process and the level of activity may multiply by several factors within the next few months due to configuration enhancements. At present the system permits job entry from the terminal, but the job must be executed in the batch mode with overnight turnaround. As soon as the configuration permits, the center will allow the option of immediate execution in batch mode and the option of printing the output on a terminal or on the printers within the Computing Center. When the turnaround gets down to a matter of minutes, the traffic should significantly increase.

In several instances, a particular administrative office has the personnel in house to develop programs in a higher level language such as FORTRAN or COBOL. The center has provided these offices with read-only capability into selected files and individuals are now programming to extract their own data. The reason for this action is primarily due to the limitations of the report writer. Even if you can accomplish a desired goal, it is inefficient when compared to a program in a higher level language, specifically written for a given task. At such time as other central offices have such capability, this tool will be extended to them.

The above three tools are already in the hands of the non-programming staff, and are being used and well received. If no canned report meets the information requirements of a user, and if the first two tools are not adequate, and the user does not have the ability to program, the request is made to the centralized programming staff. Although the Computer Center still gets many requests, the number has dropped, and staff is able to consult, design and develop the more complicated aspects of the MIS, yet not deny the user access to the information. More time is being devoted to training and documentation.

The above tools have been made available to selected offices in all instances only when a possible "need to know" encompasses all records within a given file. What about becoming a MIS to a broader community of user?

Looking into the future one can only address the capability and feasibility of such action from the computing standpoint. Whether or not the management of the University of Houston will extend the MIS further, has yet to be determined. Comments on these possibilities have been made, however no specific proposals have been presented. Again, present configuration prevents such implementation. Because this is being overcome, Center staff is now conceptualizing these extensions.

Security at the record level, and the element level becomes a more critical matter in a community MIS. The Computer Center staff has developed a workable record-level security system for its accounting system and can permit any given department, college or vice-presidential office to inquire into this system with a rather comprehensive processor already in use by the fiscal offices. This processor even crosses over into the payroll system to secure data. The entity must be registered for such processing, and must know his or her password. This being accomplished, the user may inquire into the records of his or her accounts and all those cataloged to lower echelons. This same module could be relatively easily incorporated in the appropriate packages in the user library. If this capability were extended to each college, they would be better able to monitor internal fiscal affairs.

The feasibility of interfacing this record-level security module into the report writer program is also being investigated. Initial findings have been encouraging. If such is implemented, the authorized users would be able to generate individualized management reports. Whether or not it will be feasible to package this record-security module into programs written by users has not been resolved. Complexities are greater, but more important, the dangers of circumventing this security appears quite great.

SUMMARY

A Management Information System is not a bank of information or a programmer's property, but is a tool for the management of an organization. It must be responsive to and available for management. One of the biggest challenges coming from a MIS is not how valuable the information to be retrieved may be, but how extensively one can make it availabie to a large group of users. The University of Houston is very close to having an MIS for fiscal affairs. In the coming months more capability will be available to more people. Only then will such systems mature into a Management Information System.

Chapter 11

Exporting and Importing MIS

Exportability and the In-House MIS

by Cheryl M. Traver
Stanford University

The subject of exporting an in-house Management Information System has not received much attention since most universities are more interested in importing such a vehicle. Stanford University is one of the few educational institutions which has designed and developed such a system. Called OASIS (**O**nline **A**dministrative **I**nformation **S**ystem), this data base management system was designed to serve the administrative needs of Stanford's alumni, student, personnel, and budgeting areas.

Because the decision to export the system was made too late, insufficient manpower was devoted to that end during the development phase. Consequently, the documentation was poor, and installation-dependent characteristics were not parameterized or even localized. The software was not designed for packaging, shipping, or maintenance, and flexibility in hardware was not facilitated. Much retrofitting and software modification had to be done in order to export the system.

Experience with OASIS offers a number of guidelines for those contemplating exportation. First, it is necessary to decide who comes first – the seller or the buyer. If a customer has a question, a problem, or a suggestion, someone must respond. It is difficult to place a priority on such demands, especially when a user does not promise any benefit for the home site. Secondly, one must decide on a monetary strategy. A software package will be distributed "for free", "for cost", or "for profit". The free system normally received outside funding during the development phase and therefore is distributed at the cost of installation. Unfortunately, the outside funding does not usually continue as long as the user demands. Thus, some sort of monetary agreement is normally needed to defray support costs. If no outside funding was used, a charge may be made to each new user to help defray the cost of developing the product as well as the cost of support. For software packages which are being expanded and extended, fees for updates or extensions can be charged and these "profits" then used for further enhancements to the package. Third are the distribution considerations. Does the developer want to actively market the system? Marketing is an expensive endeavor and increases the cost of developing the product. In addition, good salesmen may not have any other suitable role in the computer center.

Developing a software package to the point where it is ready for installation requires a good deal of effort. It must be a self-contained system, complete with installation instructions and provisions for training the user to tailor the software for specific applications. In addition, the system must be usable on a variety of machine configurations, some of which cannot be tested at the home site.

Updates to the distributed versions of software packages are more trouble to export than the original installation package primarily because the user wants to make updates with a minimum of effort and without harming the status of the system. Furthermore, one cannot overestimate the problems of maintaining multiple update versions for several installations. Difficulties of remembering who has which set are the least of the originator's troubles. The largest time consumer is the user who is bug hunting via the telephone. Errors are hard enough to isolate at home with a full complement of memory dumps and hardware traces, but debugging with no printed output at one's disposal and with an expensive communication device is extremely frustrating. Matters become worse if the production schedule at home begins to cause priority conflicts with outside users.

The first recommendation to anyone contemplating exportation should be DON'T. However, if one must export, one should plan for exportation during the development of the software. One must decide on a distribution strategy early and document the software well. If these steps are not taken during the development phase, it will be necessary later to devote six months and 2½ people (a software programmer, an applications-oriented training person, and a part-time editor/secretary) to produce a stable version of the software, good documentation, a method for exporting the system, and a basic training package.

It is most important to establish a set of marketing rules. One should not distribute for free. People do not trust a free product and the seller will have a constant priority struggle when outside users need help. One should not plan to gain enough money to pay for extensions desired at home. Such a serious marketing effort is risky and may not generate enough income to offset the cost of the sales effort. One should also not try to recover development costs for a software package which was developed for a single installation anyway. Instead, a software package should be provided for a charge great enough to offset the cost of realistic calculation of the manpower required for installation, training, and consultation. If possible, the seller should provide the user with a variety of cost options like a monthly maintenance fee, a one-time installation charge and a "don't call us" clause, or a fixed rate for new versions. Above all, one should keep expectations low and not expect to "sell" an MIS to anyone except oneself.

On-Line Administrative Information Systems: A Case Study

by Paul W. Sire
University of Vermont

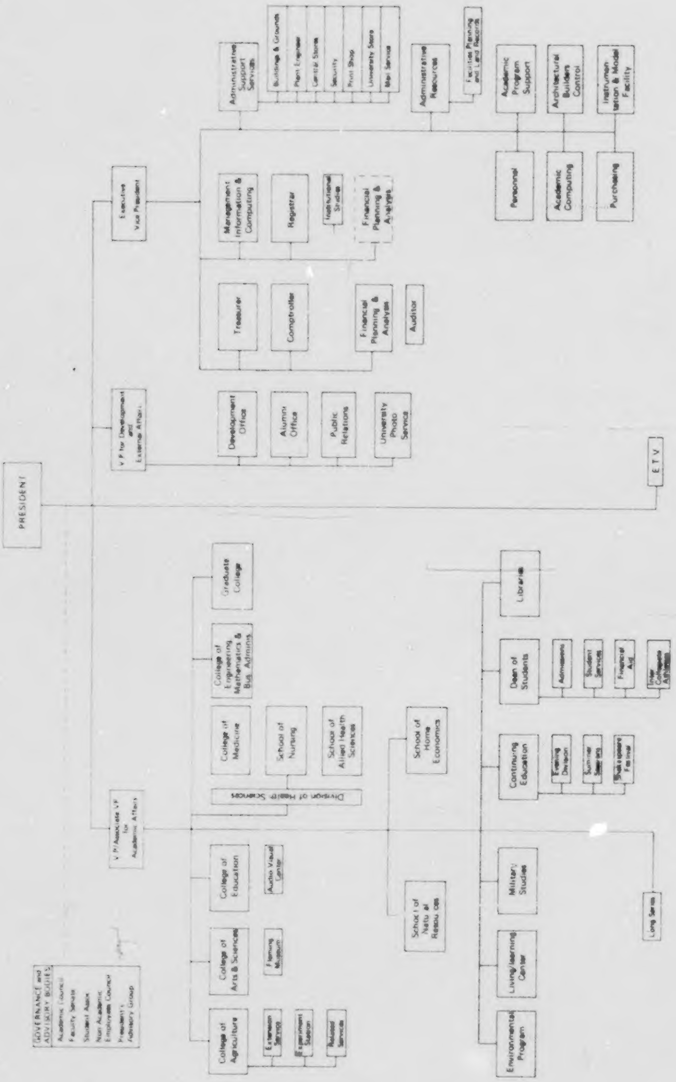
The approach taken in this paper is that of case study. Its purpose is to display administrative systems capability at the University of Vermont and what it took in time and resources to attain this capability.

The University of Vermont is a fair sized institution with an enrollment of approximately 10,000 students, full and part-time. There are two computing centers on campus, one academic and one administrative. The Academic Computation Center makes available to faculty and students time-sharing services via a Xerox Sigma 6 System. The Office of Management Information and Computing (OMIC) is the administrative computing arm of the University.

The overall structure and organization of the University is shown in Figure 11.1. All academic colleges and departments report directly to an Academic Vice-President whereas all administrative departments and offices report, at a like level, to an Executive Vice-President. The fact that the Office of Management Information and Computing reports at the Vice-Presidential level has assisted greatly in the accomplishments that have been brought to bear in the past two years and has been extremely helpful in the transition from manual to administrative computer processing.

Prior to 1971, the track record of data processing at the University of Vermont in the generic sense was at best fair to good. Data was invariably

Figure 11.1


BOARD OF TRUSTEES
PRESIDENT
GOVERNANCE and ADVISORY BODIES
Academic Council
Faculty Senate
Student Assoc
Non Academic Employees Council
President's Advisory Group
V P for Development and External Affairs
Development Office
Alumni Office
Public Relations
University Photo Service
Executive Vice President
Treasurer
Comptroller
Financial Planning & Analysis
Auditor
Management Information & Computing
Registrar
Institutional Studies
Financial Planning & Analysis
Administrative Support Services
Buildings & Grounds
Plant Engineer
Central Stores
Security
Print Shop
University Store
Mail Service
Administrative Resources
Facilities Planning and Land Records
Academic Program Support
Architectural Builders Control
Instrumentation & Model Facility
Personnel
Academic Computing
Purchasing
E T V
V P/Associate V P for Academic Affairs
College of Agriculture
Extension Service
Experiment Station
Related Services
College of Arts & Sciences
Fleming Museum
College of Education
Audio Visual Center
Division of Health Sciences
College of Medicine
School of Nursing
School of Allied Health Sciences
College of Engineering Mathematics & Bus. Admins
Graduate College
School of Natural Resources
School of Home Economics
Environmental Program
Living/learning Center
Military Studies
Continuing Education
Evening Division
Summer Session
Shakespeare Festival
Dean of Students
Admissions
Student Services
Financial Aid
Inter Collegiate Athletics
Libraries
Long Series

inaccurate, incomplete and untimely. At this point in time, one computer system, an IBM 360/44, was shared by the academic and administrative community. In the spring of 1971, it was decided that two centers, each distinct from the other, would be established. Furthermore, the administrative center would pursue the implementation of, and be dedicated to, a data management system and a data base approach to integrated files and applications development. Toward this end, IBM's IMS/GIS package was examined, but found inadequate. Limitations of the IBM 370/145 256K memory environment, inability to handle a design requirement of 10-20 CRT terminals, and a $20,000 per annum rental price rendered IBM's data management entry beyond the realm of further serious consideration. OASIS, Stanford University's Project INFO On-Line Administrative Information System, proved to be a viable alternative. Following upon an examination of OASIS services and hardware requirements in the fall of 1971, an initial version of the system was acquired in December of that year. In March 1972, the first two Sanders 720 video display terminals were installed and were made operational against a test version of a Personnel file. The ability to rapidly demonstrate OASIS services (Query and Report Writer) against an actual file went a long way toward creating a measure of acceptance and enthusiasm, on the part of administrators, of on-line files, terminals, data management software and the like.

Figure 11.2 makes reference to the staff capability of the Office. In the systems development area there are ten people who are involved heavily in maintenance as are most administrative computing groups. A major problem is maintenance of existing systems, while simultaneously creating a data base design in priority areas of concern. Priority areas at the University of Vermont have been: Personnel/Payroll/Salary, Distribution; Facilities; and Student Records encompassing Admissions and Financial Aid data. Overall, the Office has a staff of thirty: ten engaged in development; supervisory and clerical personnel; and the remainder in data center operations and data preparation. The budget of the Office is some $565,000 for fiscal year 1974 which is approximately evenly distributed between staff salaries and operating expenses.

OMIC uses an IBM 370/145 with a memory of 256K, seven disk spindles, three tape drives, and related equipment. (See Figure 11.3) The on-line data bases reside on 3.5 spindles. Currently eleven CRT terminals installed throughout the institution are located in the Registrar's Office, Financial Aid, Admissions, Personnel/Payroll and within OMIC. Growth to more than fifteen to eighteen administrative terminals on campus is projected for the foreseeable future.

OMIC computer applications are, for the most part, traditional (See Figure 11.3) and many of them existed prior to the establishment of OMIC. In Figure 11.4, those systems noted with an asterisk are on-line OASIS Systems that allow for inquiry, generation of reports and file

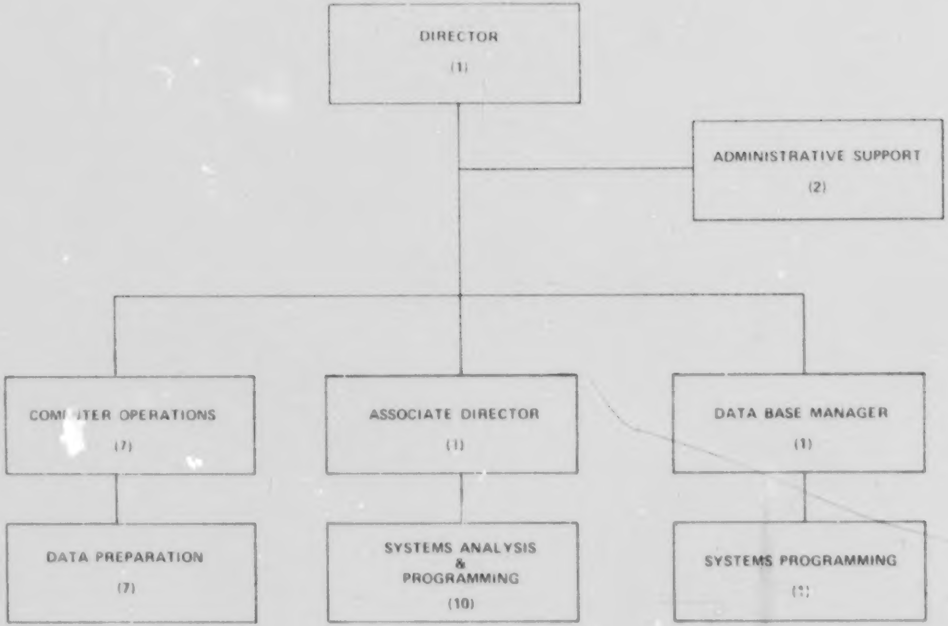


Figure 11.2 University of Vermont Office of Management Information and Computing

DIRECTOR
(1)
ADMINISTRATIVE SUPPORT
(2)
COMPUTER OPERATIONS
(7)
ASSOCIATE DIRECTOR
(1)
DATA BASE MANAGER
(1)
DATA PREPARATION
(7)
SYSTEMS ANALYSIS & PROGRAMMING
(10)
SYSTEMS PROGRAMMING
(1)

Figure 11.3 OMIC Hardware, 1973

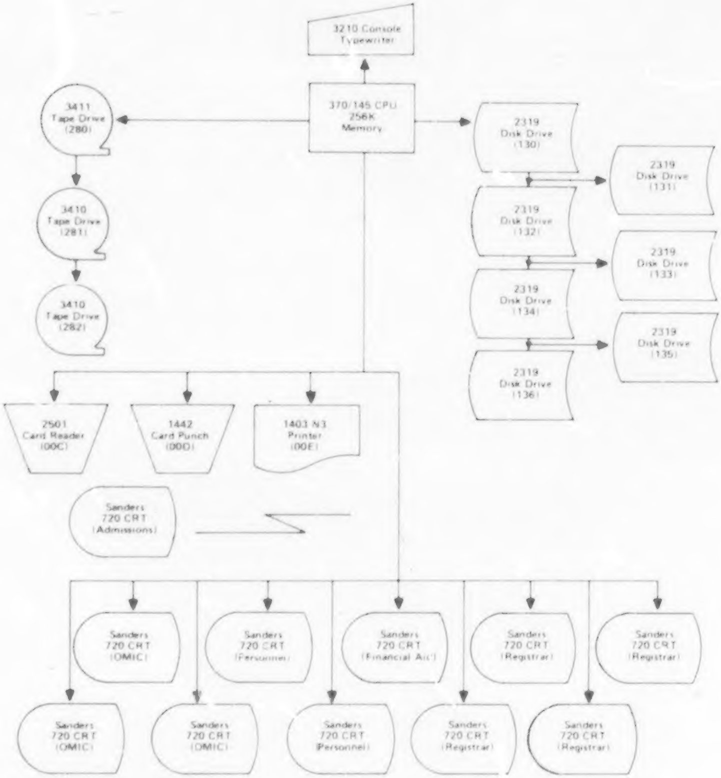

3210 Console Typewriter
370/145 CPU 256K Memory
3411 Tape Drive (280)
3410 Tape Drive (281)
3410 Tape Drive (282)
2319 Disk Drive (130)
2319 Disk Drive (131)
2319 Disk Drive (132)
2319 Disk Drive (133)
2319 Disk Drive (134)
2319 Disk Drive (135)
2319 Disk Drive (136)
2501 Card Reader (00C)
1442 Card Punch (00D)
1403 N3 Printer (00E)
Sanders 720 CRT (Admissions)
Sanders 720 CRT (OMIC)
Sanders 720 CRT (Personnel)
Sanders 720 CRT (Financial Aid)
Sanders 720 CRT (Registrar)
Sanders 720 CRT (Registrar)
Sanders 720 CRT (OMIC)
Sanders 720 CRT (OMIC)
Sanders 720 CRT (Personnel)
Sanders 720 CRT (Registrar)
Sanders 720 CRT (Registrar)

Figure 11.4 University of Vermont Systems 1973
Office of Management Information and Computing

CURRENT SYSTEMS

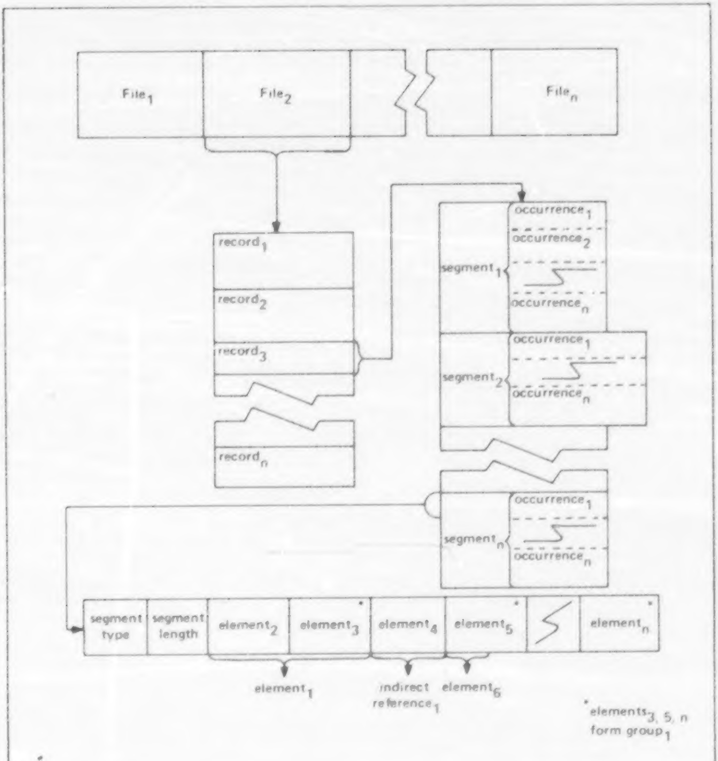
a. Accounts Payable
b. Accounts Receivable
c. Alumni Records
d. Audiovisual
e. Automobile Registration
*f. Board (College) Scores
g. Bookstore Inventory
h. Budget Planning and Analysis
i. Centrex (telephone) Billing
j. Class Scheduling
k. Computer Billing
l. Current Payroll Systems
m. Current Student System
n Data-Text System (Sociology)
o. Experimental Programs
*p. Facilities
q. Financial Accounting
r. Gift Accounting
s. Graduate College Admissions
*t. Grants Information
u. Health Sciences
v. High School Scheduling
w. Induced Course Load Matrix (ICLM)
x. Library Periodicals
y. Mail Service List
z. Medical Books
aa. Moveable Equipment
bb. Shop Stores Inventory
cc. Student Pre-Billing
dd. Tape Library
ee. Label Utilities

NEW SYSTEMS

*a. Personnel/Payroll/Distribution
*b. New Student System
    1. Admissions
    2. Financial Aid
    3. Registration
    4. Drop/Add
    5. Billing
    6. Grade Reporting
    7. Course/Section Master
    8. Counseling and Testing

*OASIS SYSTEM

Figure 11.5 OASIS File Structure


File$_1$
File$_2$
File$_n$
record$_1$
record$_2$
record$_3$
record$_n$
segment$_1$
occurrence$_1$
occurrence$_2$
occurrence$_n$
segment$_2$
occurrence$_1$
occurrence$_n$
segment$_n$
occurrence$_1$
occurrence$_n$
segment type
segment length
element$_2$
element$_3$*
element$_4$
element$_5$*
element$_n$*
element$_1$
indirect reference$_1$
element$_6$
*elements$_{3, 5, n}$ form group$_1$

updating via terminals. The systems listed include 550 computer programs.

Inquiry is accomplished through a data management OASIS service called QUERY. This service need not be predetermined but is open-ended. It allows for response to questions on the spot, on demand, as formulated at the moment. There is essentially no constraint as to what may be retrieved. Some examples of QUERY are offered in Appendix D. Under these circumstances, it might be said that OMIC has provided distributed data through a centralized computation center.

Figure 11.5 illustrates the OASIS file structure. The OASIS file is made up of traditional records which are composed of segments. Segments can occur from zero to many times, so if a piece of the data is not required at a particular point in time, it is not resident on the system. Because segmentation permits one to modify the file fairly easily, records within the OASIS file vary in length. As experience is gained with the sundry data bases, OMIC staff have found that some of the elements which are not used, are no longer required and other elements must be defined. It is a fairly simple task to modify the record and OMIC staff have already done this numerous times. The data bases have been changed many times over since their initial definition and continue to be changed. Appendix C, which is intended for the readers' general perusal, illustrates the magnitude of those primary and secondary files resident on 3.5 spindles. Appendix D illustrates a sample query and computer response using OMIC files.

In summary, on-line files development at the University of Vermont has occurred only over the past two years. This work has included basic data definitions, programming, testing, training, including all required aspects leading to systems implementation.

Importing MIS Components: A Practitioner-Manager Perspective

by Henry G. Vaughan
Cornell University

To most administrators or computer users in higher education MIS (or Management Information Systems) is a bad term. The reason is two-fold. First, the meaning of MIS in higher education is particularly obscure and varies from institution to institution depending upon whether one is talking to top management or to a technician. Secondly, and more seriously, to many less informed members of the academic community, the term MIS implies the imposition of a dictatorial type of government on the institution. For these reasons, and in the hope that a rational discussion can be possible, this paper will discuss "Information Systems" rather than "Management Information Systems" because the difference is more than semantics in most academic communities.

For purposes of simplifying the discussion to a reasonable length, an initial assumption must be made that there is agreement with, or at least toleration of the decision to acquire an information system for the university. Even with this assumption, it is worth noting that the acquisition is viewed differently by various members of the university community. Essentially the user is facing the classic black box situation. As shown in Figure 11.6, typically a "computing utility" operates a "black box" in support of the data processing needs of the institution. At larger institutions the "black box" is in reality a complex mix of sophisticated computer hardware and accompanying software operated by a large staff

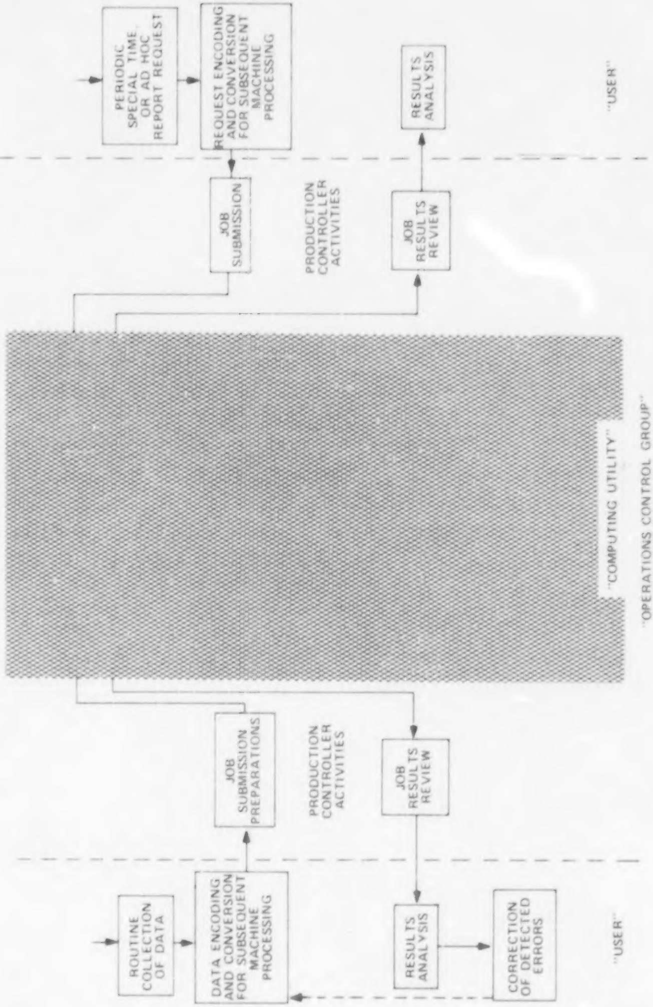

ROUTINE COLLECTION OF DATA
DATA ENCODING AND CONVERSION FOR SUBSEQUENT MACHINE PROCESSING
JOB SUBMISSION PREPARATIONS
PRODUCTION CONTROLLER ACTIVITIES
JOB RESULTS REVIEW
RESULTS ANALYSIS
CORRECTION OF DETECTED ERRORS
"USER"
"COMPUTING UTILITY"
"OPERATIONS CONTROL GROUP"
JOB SUBMISSION
PRODUCTION CONTROLLER ACTIVITIES
JOB RESULTS REVIEW
PERIODIC SPECIAL TIME OR AD HOC REPORT REQUEST
REQUEST ENCODING AND CONVERSION FOR SUBSEQUENT MACHINE PROCESSING
RESULTS ANALYSIS
"USER"

Figure 11.6 The Black Box

of "technicians". The users are external to the "black box" and are often even one step further removed. Historically a buffer group, the Operations Control Group has existed between the users and the "black box". People in the Group called Production Controllers have had the very important function of serving as a translator that instructs the computer on behalf of novice users and obtain results for them. This functional departure point is now changing quite drastically due to changes in technology developed at the insistence of others. One of the questions that must be asked is "Will the traditional mode exist in the future?"

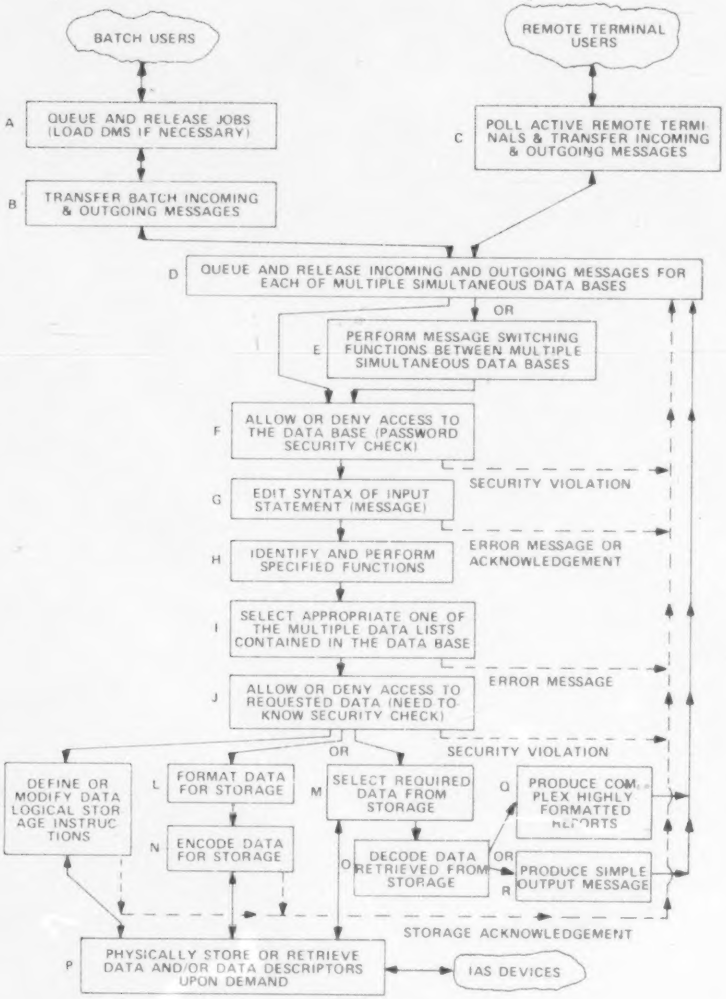
In order to design the Information System and define its contents, it is necessary to know how the user wants to interface with the black box (i.e., what is the desired mode of user interaction?). Is the institution going to assign Production Controllers as intermediaries or is the experienced user expected to directly interact with the system through a terminal device? What kind of response considerations is wanted: instantaneous response or a response time of several days? Does one want information that is an abstraction of the data designed to convey meaning, or does one want a pretty report? Information can of course be obtained faster than a pretty report. The answer, of course, depends upon the sophistication of the user. Are integrated files in that black box necessary or not? There are certain advantages of integrated files but there are also disadvantages including the time and expense associated with constructing the system. Is a separate Data Base Management System (DBMS) necessary as a distinct piece of software, or should that function be defined through the way application solutions are coded?

If a university decides to import a data management system package, it must define before importation the desired characteristics and sequence of operations desired and appropriate to the University's normal *modus operandi*. One possible technical description of appropriate characteristics is shown in Figure 11.7. Although the subject of data base management systems is very complex technically, it should be emphasized that no standards exist in this area. Some Data Base Management Systems do certain tasks and others perform somewhat different roles. Whether a DBMS should be oriented to batch or teleprocessing is only one of the characteristics that must be defined by the institution. Many of the functions shown can and should be done by the host computer's operating system although early data base management systems did not do it that way.

DATA BASE MANAGEMENT SYSTEMS

Three distinctly different types of Data Base Management Systems seem to exist. The first group, file management systems, is characterized by software such as the Informatics' original MARK IV package. These

Figure 11.7 Simplified Total Functional Flow Necessary for a DBMS Application


BATCH USERS
REMOTE TERMINAL USERS
A
QUEUE AND RELEASE JOBS (LOAD DMS IF NECESSARY)
B
TRANSFER BATCH INCOMING & OUTGOING MESSAGES
C
POLL ACTIVE REMOTE TERMINALS & TRANSFER INCOMING & OUTGOING MESSAGES
D
QUEUE AND RELEASE INCOMING AND OUTGOING MESSAGES FOR EACH OF MULTIPLE SIMULTANEOUS DATA BASES
OR
E
PERFORM MESSAGE SWITCHING FUNCTIONS BETWEEN MULTIPLE SIMULTANEOUS DATA BASES
F
ALLOW OR DENY ACCESS TO THE DATA BASE (PASSWORD SECURITY CHECK)
SECURITY VIOLATION
G
EDIT SYNTAX OF INPUT STATEMENT (MESSAGE)
ERROR MESSAGE OR ACKNOWLEDGEMENT
H
IDENTIFY AND PERFORM SPECIFIED FUNCTIONS
I
SELECT APPROPRIATE ONE OF THE MULTIPLE DATA LISTS CONTAINED IN THE DATA BASE
ERROR MESSAGE
J
ALLOW OR DENY ACCESS TO REQUESTED DATA (NEED-TO-KNOW SECURITY CHECK)
OR
SECURITY VIOLATION
K
DEFINE OR MODIFY DATA LOGICAL STORAGE INSTRUCTIONS
L
FORMAT DATA FOR STORAGE
M
SELECT REQUIRED DATA FROM STORAGE
Q
PRODUCE COMPLEX HIGHLY FORMATTED REPORTS
N
ENCODE DATA FOR STORAGE
O
DECODE DATA RETRIEVED FROM STORAGE
OR
R
PRODUCE SIMPLE OUTPUT MESSAGE
STORAGE ACKNOWLEDGEMENT
P
PHYSICALLY STORE OR RETRIEVE DATA AND/OR DATA DESCRIPTORS UPON DEMAND
IAS DEVICES

packages perform only a limited subset of the appropriate cited functions. Some degree of security is provided through password protection schemes for restricting access to the data base, simple arithmetic calculations can be directly performed, and a fairly extensive report writer usually exists.

A second type of Data Base Management System is that class best described as user language systems. Examples are IBM's GIS or TRW's GIM system. Again, only a limited subset of the appropriate cited functions are performed, but the emphasis is on a different set of characteristics. Typically the user of such a system interacts, often through teleprocessing, with a system that translates his or her English-like statements into computer-processable code. Sophisticated application programs and high-level languages such as COBOL, PL/1 or FORTRAN are not typically used or permitted.

The third type of generic Data Base Management System are those best characterized as an extension of the host computer's operating system. Examples are IBM's IMS and the "Exec" portion of Stanford University's OASIS. These systems similarly do not attack the total scenario of appropriate cited functions previously described, but limit themselves to handling the very complex technical subjects such as message switching, file definition, and data base accessing. Omitted functions include selection criteria for reports, definition of report formats, and some of the more complex arithmetic calculations. Special user-prepared applications programs must be written to cover the functions omitted.

INSTITUTIONAL SELF-STUDY

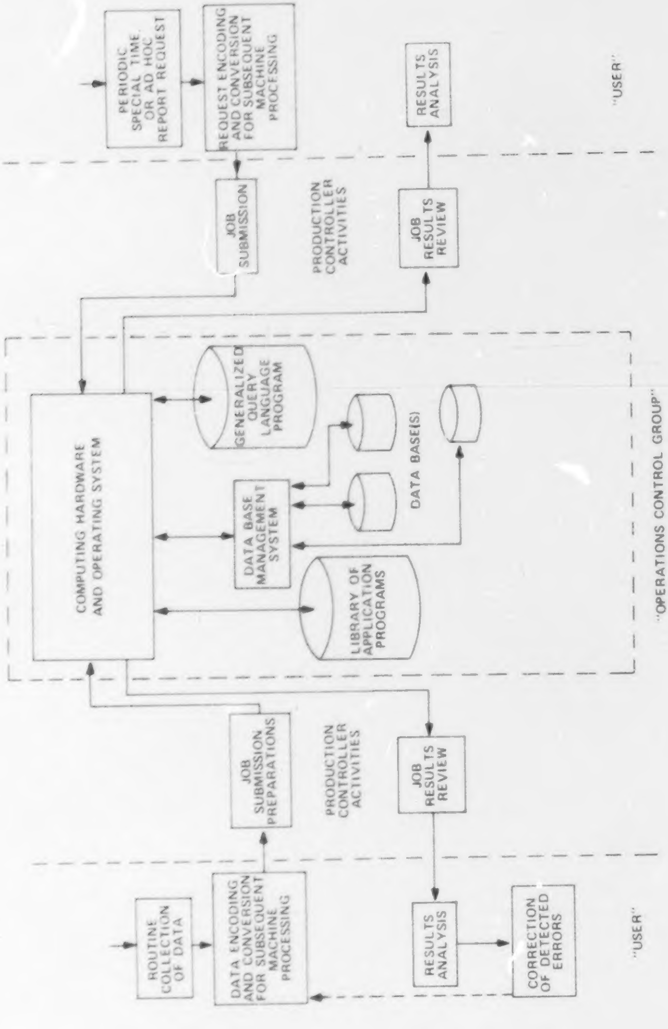
An institution that is going to import a Data Base Management System must define a desired scenario of operation before a system can be selected. One type of system will manipulate simple files for novice batch users while another will provide complex file manipulation capabilities to users who know how to use sophisticated data bases and who want to access them from remote terminals. In the latter instance, a system oriented to the efficient processing of a large volume of transactions would be required. In such instances very little use is made of Boolean-logic file searches for management information, or of report formatting capabilities that the DBMS might possess. In contrast, a small college without its own computer, but with access to a nearby computer, may want an administrative data processing environment with essentially the opposite characteristics. Thus, one of the largest administrative decisions facing institutions of higher education is to determine the general characteristics they individually want in an imported Data Base Management System. That single answer will determine, in large part, the responsiveness and efficiency of data processing operations, and to an increasing degree, the characteristics of the institution's administration.

As shown in Figure 11.8, the inside of the previously described "black box" is complex and often chaotic, but an ordering is rapidly occurring as the use of Data Base Management Systems become more widespread. In the future the information system software components of the larger universities will be organized as shown in that chart. First, although all accesses to the computing hardware will continue to occur through the operating system, the interaction by users will be more direct and not through intermediaries. The data base management system will manage the input of data into the data bases and determines how it is retrieved. A vast library of application programs will exist which will match and be tailored to the individual institution's functional data processing requirements. A generalized query langauge program will also be widely used to supplement the other software packages and to provide better capabilities for Boolean-logic file searches and report-formatting. At Cornell University, IBM's Version 2 is currently used for data base management and Informatics' MARK IV package is currently used for the generalized query language program. Since administrative data processing is in transition at Cornell, as elsewhere, the MARK IV package (with an IMS interface) is used to process against both the IMS data bases, where they exist, or against regular files. Stanford and some other institutions are using OASIS. In that case a Query Module, which is a supplied application program handled under the data base management system, serves as the generalized query language program. Conceptually, these administrative data processing environments are very similar.

THE NEED FOR SELF-STUDY

Before embarking on the acquisition of an information system, the institution should do a detailed self-study of its needs, environment, and finances. A crucial part of the study should be an overall conceptual plan defining the objective of the undertaking. In this plan it is necessary to be quite specific. One must decide, for example, whether it is desirable and practical to integrate files of records on students, applicants and alumni. At a large institution like Cornell University, it becomes prohibitively expensive to process all of that data in one integrated file. Therefore, separate data bases are typically used but which are capable of being linked, if desired for special purposes, through a common element such as a Social Security Number. Many times data processing application systems must undergo extensive revisions to approach the ideal of the institution's plan. For example, at Cornell, the old student records system uses a unique student identifier number in lieu of the Social Security Number. In contrast, the new system that we are developing will use the common linking element of Social Security Number in order to enable better cross-matching of data elements in several separate data bases when needed.

Figure 11.8 Inside the Black Box


ROUTINE COLLECTION OF DATA
DATA ENCODING AND CONVERSION FOR SUBSEQUENT MACHINE PROCESSING
RESULTS ANALYSIS
CORRECTION OF DETECTED ERRORS
"USER"
JOB SUBMISSION PREPARATIONS
PRODUCTION CONTROLLER ACTIVITIES
JOB RESULTS REVIEW
COMPUTING HARDWARE AND OPERATING SYSTEM
LIBRARY OF APPLICATION PROGRAMS
DATA BASE MANAGEMENT SYSTEM
GENERALIZED QUERY LANGUAGE PROGRAM
DATA BASE(S)
"OPERATIONS CONTROL GROUP"
JOB SUBMISSION
PRODUCTION CONTROLLER ACTIVITIES
JOB RESULTS REVIEW
PERIODIC, SPECIAL TIME, OR AD HOC REPORT REQUEST
REQUEST ENCODING AND CONVERSION FOR SUBSEQUENT MACHINE PROCESSING
RESULTS ANALYSIS
"USER"

For each component of the institution's information system, the institution must define the specific requirements that must be met. The appropriate statement of requirements must also specify any constraints on solutions. For example, one of the fundamental constraints is usually money. How many dollars can the institution risk to lease or purchase a piece of software? This very fundamental first consideration may eliminate half of the alternatives. What, if any, is the role of the Data Base Management System? Is the desired solution to do very complex things, or something very simplistic? Is the application to be transaction-oriented, batch-oriented, teleprocessing-oriented, or management Boolean-logic file search oriented? Most importantly, what are the user's office functions? Is the system being acquired for the personnel administration function, for the Payroll Office, or for an integrated payroll/personnel system? Is the system to be used primarily for registering students, or to provide information on students directly to the registrar and the academic deans? The answers are important since the data processing solution will only be as good as the definition of the problem that undergirds the design of the solution.

THE BIDDING PACKAGE

If one continues this sort of self analysis of its information system component needs to what may be an extreme, the importing institution will prepare a specific set of formal requirements against which candidate solutions can be evaluated. The bidding package is a formal statement of the problem and of constraints upon solutions that would be useful, if appropriate, to soliciting competitive bids to procure the software solution.

Cornell recently prepared a bidding package in conjunction with its decision to import someone else's student records application system. The package included a description of the institution in general terms with relevant specific detail, and the citation of appropriate general policies such as the expectation of a continued orientation toward full-time students. Other institutional policies regarding the roles of the various component colleges and centralization of authority over aspects of particular functions were also stated. Another important section of the package dealt with the current and expected computing environment of the institution. The university has a centralized computing utility with certain characteristics and with certain established policies that affect any computer program operating on the facility (e.g., maximum core size or maximum number of permitted set-up devices). Specific administrative computing constraints were also defined such as the requirement that a previously-selected Data Base Management System be used for certain roles. The source language desired for application programs was also specified.

The bidding package also has to present the results of a functional analysis of the tasks to be done by the system. Detailed specific user requirements like editing lists, codes or translations shown in reports, totals, and the needs for inverted file accessed all should be defined.

Subsequent to completing the self-study and preparing the formal statement of requirements for the application, alternatives should be proposed in writing. In Cornell's search for a student records system to import, three alternatives were evaluated. Two proposal/study reports were prepared by commercial vendors and a third was prepared by the Cornell data processing staff in conjunction with the staff of the potentially-exporting institution.

THE IMPORTING PROCESS

Although importing application software from a similar institution looks very attractive financially, the old axiom is still true that nothing is ever free. Even though the software may come to the importing institution without charge, or for a minimum fee, there are many other associated costs. When Cornell University imported a student records system from a commercial vendor the costs of the software obtained, on a full-cost basis, were only about one-third of the total monies needed. Other significant costs were incurred for data conversion, user training, additional staff, forms, and minor facility modifications.

The criteria of prime importance to the importing institution are generally institution-unique. Although the general considerations may pale in comparison, they should not be ignored. One of the more significant of these is the maintenance support offered by the exporter. Others include: available documentation and planned documentation maintenance; the level of assistance in implementation offered by the exporter including user training; explanation offered by the exporter of the source language and source code in which the software is written; and some idea of the perceived efficiency of the software.

After all of the alternatives have been evaluated and a selection made that can be justified to the world including the faculty, one should establish a project team and charge it with responsibility for completing importation of the information system component. The relationship between the project team, the university data processing organization and others should be formally defined to minimize future problems. Finally, the management of the importing institution must specify the acceptance criteria for the system and the procedures for both testing it and using it.

It is doubtful that an information system has ever been imported successfully without problems. One must expect the unexpected and also allow time to fix things. By taking someone else's system, an institution is implicitly deciding to accept a system that, by definition, will never exactly meet its needs. The importing institution should, therefore, allow time to make some changes to better match the information system component to the new environment. However, moderation should be observed as the effort to make changes may exceed the cost of developing a new and more appropriate system. It is also appropriate to *not* modify non-application software, unless one has very sophisticated needs, a very sophisticated staff, and adequate financial resources. Because few colleges and universities can afford the luxury of a custom designed administrative information system, in the future, more can be expected to join the ranks of administrative software importing institutions.

One University's View of Data Management Systems

by Kenneth E. Shostack
Harvard University

Because of the current financial situation, many colleges and universities are re-examining their approach to administrative data processing. One concept most frequently offered for improving services is that of data management systems, sometimes called data base management systems. At Harvard University the Office of Information Technology was asked whether a data management system would alleviate the three problem areas of maintenance, timely reporting, and data entry; and which data management system would best contribute to the alleviation of those problems. In order to respond to these questions it was necessary to define problems specific to Harvard University and study the available data management systems. This paper describes some of the problems faced by Harvard and discusses those aspects of the evaluation of available data management systems which are not normally described in the literature.

Harvard University is facing three problems which are being faced by most data processing departments. The first of these problems is application system maintenance. As shown by several recent studies, costs associated with data processing have been shifting. In the early 1960's the most costly item in the data processing budget was computer resources. By the mid 1960's the most costly item in the budget was applications development. Today, the most costly item in the budget is maintenance including adding, deleting or modifying of data elements in a file, i.e., changes which usually require modifying every program in the system. One recent study[1] claimed that over 50 percent of the data processing budget in a mature operations organization is for maintenance. Another 30 to 40 percent of the budget is for operations. This leaves only 10 to 20 percent of the budget for developing new applications.

A second problem is preparing timely reports, especially new reports or one time reports. Historically, requests for special reports were fulfilled by developing new programs and usually the need for the reports ceased to exist before the report was ready. More recently report writers have been available which enable programmers or analysts in user departments to prepare reports in a matter of hours or days rather than weeks. Users now want more timely information, such as that available with on-line terminals and query languages.

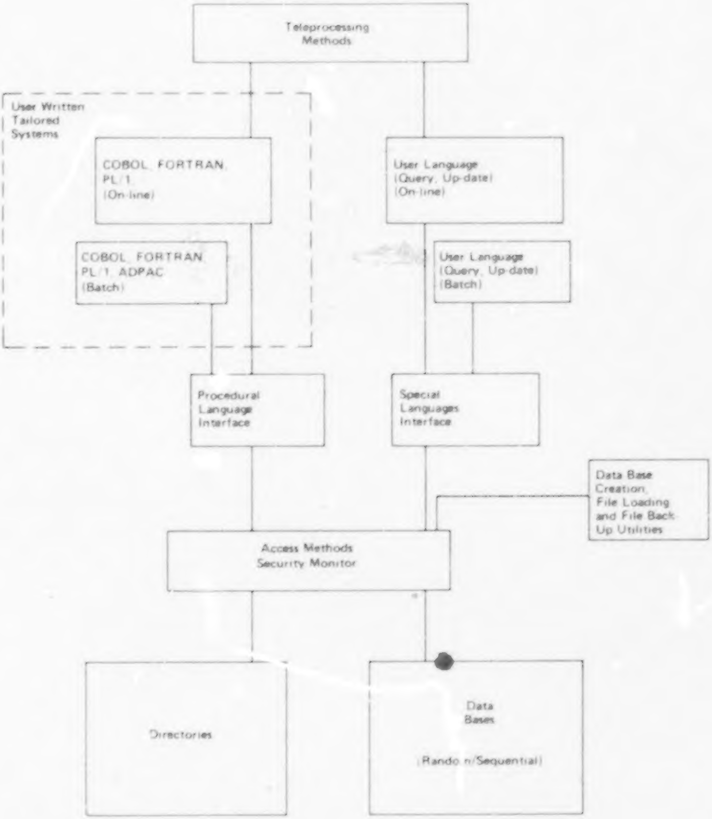
A third problem is getting transaction data into the data files or data base in an efficient and timely manner. There are many parameters to measure whether a method of data entry is efficient, and there are many philosophies regarding approaches to data entry. Currently, all administrative data entry at Harvard is accomplished using keypunches or a key to disk system. The cost effectiveness of alternative data entry needs to be explored.

FRAMEWORK FOR ANALYSIS

In order to assist in the discussion of the requirements for a data management system and the capabilities offered by various vendors, a framework for analysis was developed. This framework is based on a functional definition of the ideal or complete data management system. The complete data management system (See Figure 11.9) is a set of computer programs which provides access to an aggregate of data elements called a data base.[2] The data base management system provides a method of gathering data elements, eliminating duplication of information, reducing file space, and achieving greater accuracy (by reducing redundancy).

The data is normally stored on a disk in a random manner which provides for: 1) sequential access for volume processing; and 2) random access based on multiple access points (inverted value tables) to facilitate random retrieval and maintenance. Security of confidential information is maintained at the data element level as well as at the data base level. Services are provided for generalized and tailored access to the data base in on-line mode via a teleprocessing monitor from various typewriter and video terminals or batch mode for high volumes of transactions. The generalized services include data base inquiry, data base update, and the quick specification and generation of reports through the user language interpreter. The tailored services are user written application programs,

Figure 11.9 Complete Data Management System


Teleprocessing Methods
User Written Tailored Systems
COBOL, FORTRAN, PL/1, (On-line)
COBOL, FORTRAN, PL/1, ADPAC (Batch)
User Language (Query, Up-date) (On-line)
User Language (Query, Up-date) (Batch)
Procedural Language Interface
Special Languages Interface
Data Base Creation, File Loading and File Back Up Utilities
Access Methods Security Monitor
Directories
Data Bases
(Random/Sequential)

created to perform functions through procedural language interfaces.

The central facility of a data base management system is the data access program which accepts requests from higher level programs and translates them into detailed manipulations of data elements in the data base. Requests can refer to data elements by name, without regard to their physical location within the data base. Thus, it is possible to deal with the data in a logical manner without regard to data base organization.

THE HARVARD TESTS

This framework can be used to discuss the four products which the Harvard Office of Information Technology examined in depth. (See Figure 11.10)

Figure 11.10 Administrative Systems Examined by Harvard

Product	Vendor
IMS, GIS, IQF	IBM
MODEL 204, IFAM	Computer Corporation of America, Cambridge, Mass.
System 2000	MRI Systems Corp., Austin Texas
TOTAL, Environ/1	Cincom Systems, Inc. Cincinnati, Ohio

A fifth system, OASIS, which was developed by Stanford University, was also considered. OASIS is an excellent system, but it does not now offer the flexibility required to fit into Harvard's plans. Any college or university considering a data management system should give OASIS some consideration.

As a result of an examination of these systems, several general observations were made. First, no system offered all the features desired by Harvard University. Second, except for operating systems, data management systems are the most complex software systems available. Few installations choose an operating system independent of the hardware. Third, the monthly rental for a complete package from each of the four vendors ranged between $3,000 and $3,400. The monthly rental for the features which Harvard would acquire ranged between $1,700 and

$2,400. Fourth, the introduction of data management systems appear to have the following effect on applications systems costs:

Development Costs	
Procedural Language Applications	20% Lower
Special Language Applications	60% Lower
Operating Costs	20% Higher
Maintaining Costs	20 to 40% Lower

Thus the costs are shifting from the labor intensive processes to machine intensive processes. Finally, by eliminating data redundancy, the department which maintains the file would have the most efficient access to that file. Other departments would have less efficient access, but would be able to eliminate the clerical tasks required to maintain the data.

Recognizing the large differences in capabilities of these systems and the variety of techniques used to implement any given function, it was decided that the best method for obtaining an understanding of each system would be obtained by running some tests using data from existing systems. The tests were not intended to be complete tests of each system, nor were they intended to show all the strengths and weaknesses of each system. They were intended to provide a deeper understanding of how these systems could be used, and to augment the conversations with users, and the study comparing the features of each system. Rather than documenting a matrix of features and a description of conversations with users of each system, the reader is referred to the documents complied by the CAUSE Data Management Systems Task Group which are available from CAUSE and the reports of the CODASYL Data Base Task Group which are available from the Association for Computing Machinery.

Harvard University currently maintains two personnel record systems, one for salary and wage personnel and a second for employees who hold Corporation appointments. The initial application schedule for implementation using a data management system is an Employee Information System, which will be an upgrading and integration of the two separate personnel systems. A. M. Koss, a Senior Staff Analyst in the Applications Development Group, was the project leader responsible for developing the wage and salary personnel system, and is the project leader on the new Employee Information System. She designed the tests using data from the two existing systems. These tests included loading 500 records and performing 53 transactions against the data. The transactions included: ten insertions; seven deletions; seven replacements; and twenty-nine retrievals. Costs of loading the data are shown in Figure 11.11 and transaction

Figure 11.11 500 Personnel Records Test Data

Load	Core	CPU Time	Disk Accesses	Costs 370/165 OS/ASP	Disk Storage (Bytes)
IMS	162K	.087 min.	283	$2.27	117K (1 Key)
MODEL 204/IFAM	200K	.216 min.	2791	$5.26	173K (5 Keys)
SYSTEM 2000	146K	.198 min.	721	$4.19	139K (5 Keys)
	314K	.215 min.	607	$5.23	
TOTAL	70K	.118 min.	2877	$6.87	179K (5 Linkage Paths)

Figure 11.12 Performance for 53 Transactions

	Core	Disk Accesses	CPU Time	Cost	Programming Cost
IMS	150K	315	.015 min.	$1.04	$600
MODEL 204/IFAM	200K	622	.118 min.	$4.03	$240
SYSTEM 2000	138K	4551	.101 min.	$11.18	$240
	306K	477	.035 min.	$2.28	
TOTAL	70K	620	.079 min.	$2.85	$600
Transactions		Insert Data – 10 Delete Data – 7 Replace Data – 7 Retrieve Data – 29			

Figure 11.13 Additional Test Data

1. COBOL Program (Batch)

	Core	CPU Time	Disk Accesses	Cost
MODEL 204/IFAM	200K	.081 min.	985	$3.63
SYSTEM 2000	170K	.079 min.	5540	$12.59

2. User Language On-line

	CPU Cost	Core Cost	I/O Cost	Connect Cost	Total
MODEL 204/IFAM	$1.30	$5.74	$2.33	$1.19	$10.57
SYSTEM 2000	$1.28	$7.16	$6.20	$1.94	$16.58

performance for processing transactions is shown in Figure 11.12. For simplicity, the charts do not show all components of the cost, but only the more common components of cost, namely: cpu, and disk accesses. By using program overlays, SYSTEM 2000 can be configured to run in various size regions of core. The test run at Harvard (See Figures 11.11 and 11.12) were run in the smallest (136K) and largest regions (314K). Performance differences are the most pronounced on the tests of transaction performance (Figure 11.12). One item of interest from the tests was the ability to develop applications programs in 60% less time with the special languages provided by SYSTEM 2000 and Model 204, than with COBOL as required for IMS and TOTAL.

In addition to the tests described above, two additional tests were performed with SYSTEM 2000 and Model 204/IFAM. These two tests (See Figure 11.13) involved the fifty-three transactions with a batch COBOL and also with the user language under TSO in an interactive mode. In both tests the highly overlaid version of SYSTEM 2000 was used.

The Harvard Computing Center operates an IBM 370/145 which is connected via microwave link to an IBM 370/165 at MIT. The IBM 370/165 has Time Sharing Option (TSO) and low speed telecommunication access ports available to users. An analysis of the costs for dedicating a region and low speed telecommunication ports on the IBM 370/145 to data management system and the costs of operating a data management system under TSO, showed that until there were six terminals on-line three hours per day it would be more cost effective to operate in the TSO environment. One disadvantage might arise because each user would have his own copy of the data management system software. There would be a potential problem of two users attempting simultaneously to update the same record. A further disadvantage of the TSO implementation is that users are normally restricted to regions of less than 200K. Because solutions are available for this potential problem, it was decided that, initially, TSO would be the teleprocessing monitor. As a result the Harvard University Office of Information Technology determined that MODEL 204/IFAM would provide the data management functions required by the University. This system will be used to develop the new Employee Information System and other applications. After each system is implemented post-implementation audits are planned to determine whether MODEL 204/IFAM is performing as expected.

REFERENCES

1. "First Warned to Design Systems for Small, Unsophisticated User", *Computerworld* Vol. VII, No. 40, October 3, 1973 Pg. 27.

2. This description is partially extracted from "OASIS, Briefly Stated" in the OASIS Programmers Reference Booklet. (Project Info Stanford University, Stanford, California) Lyle Settle, Editor

3. Other members of the Office of Information Technology staff who worked on this project include D. J. Farrell, A. M. Koss, E. Ekstrand, J. Noonan, G. Thurow, A. Paxson and R. Carroll.

PART IV

COMPUTING
IN
INSTRUCTION

185/186

Chapter 12

Computing in Instruction: An Overview

by William Atchison
The University of Maryland

An overview of computers in instruction could have as a subtitle, Computing in Transition. The use of computers in instruction has been changing rapidly as more and more use of computers has been made. The Fall 1973 EDUCOM Conference included far more indications of the use of computers for instruction than many previous meetings.

The kind of computer-based activities for instruction evident at the University of Texas at Austin including CAI, CMI, simulation and class record keeping are indicative of activities at many schools. At the University of California, San Diego, 7,000 students are using the computer for 200 courses, primarily in batch processing mode.

Computing for instruction is in transition but no one alternative is dominant. Large, general purpose, single facility computer centers may be ending for many educational research institutions but the computer is growing in importance in a university like Texas. The University of Maryland is also certainly at a transition stage. Martin Greenberger, in an editorial in *Science*, recently hedged his bets a little bit, indicating some big university computer centers will survive. The University of Maryland is moving a little more strongly than some other schools toward growth based on instructional use. In 1973, 70% of computer jobs are for instruction. This may also be happening in a great many places. Faculty at Maryland are talking about now doing some of the things that Texas is already doing: for example, having a large screen in front of the classroom and projecting some of the computer output. New remote terminals are also being requested on the several University of Maryland campuses. One does hope that some of the current work at the University of Texas Project C-BE and CONDUIT will do something to help break the "Anastasio" cycle. If the cycle can be broken, additional funding for computing in instruction may also become available. The balance between batch processing and interactive computing at university centers may also change as institutions provide less computing for research. Gerry Weeg at the University of Iowa is convinced at this stage that he must move toward timesharing to take care of his students. Dartmouth has traditionally emphasized timesharing service for instruction. On the other hand, Ken Bowles at the University of California, San Diego, relies primarily on batch processing to serve instructional needs at UCSD.

In some cases, significant applications of computing for instruction have been demonstrated. Particularly impressive is Gene Geisler's report on instruction-oriented computing in the California State University and Colleges system. Nineteen semi-autonomous campuses with 280,000 students and 13,000 faculty utilize a regional network for instruction. From the start, computers in instruction had been emphasized within that California system. Good people in practically every departmental area in many of the 19 universities have worked together to develop the regional network. In the California computer network, each school can take advantage of what's happening in each of the others whether over one of the shorter communication links or over the largest link which is 700 miles. A great deal of creative work has been done in the California system. Workshops for faculty proved to be very successful in developing instructional use of the resources on the California State University and Colleges Network.

Keith Hall and others from Pennsylvania State University have been working in the area of computer-aided instruction quite strongly. Two large vans which have been set up with a computer and 16 terminals as well as communications couplers to allow access to Penn State computing by traveling to the schools to instruct some of the teachers in the various areas in the state. The Pennsylvania State University mobile learning center represents a different idea that is growing in popularity for the area of special education as well as traditional in-service education.

Under the direction of Pranas Zunde at Georgia Institute of Technology, an audio-graphic learning facility has been developed in which a computer is installed in one facility where one can listen, do some teaching and talking about CAI packages, or simply use the packages. One reasonably large research project underway at the AGLF is a study of the transmittal of scientific information. Much research work that has been done in various research agencies throughout the country has never really been fully used by other scholars. The study seeks ways to make such research more easily available using the computer.

A group at the National Bureau of Standards, with NSF support, has begun work on criteria for evaluation of the performance of some of the existing computer networks. This, of course, is a very tough problem. Some of the items under study are documentation standards and the possibility of using "network access computers" to ease the process of getting into a network for the user.

Two well-known projects in computer-assisted instruction, the PLATO and TICCIT systems, are now being evaluated by a team of researchers led by Ernest Anastasio of the Educational Testing Service. The evaluation seeks to answer questions like the following: What are the replication costs? What are the usage costs? How much would it cost for a given college to move into these things? What is the learning effectiveness of the CAI system? Is the system accepted by users? Is it reliable? Is it easy to use? These and other questions must be addressed if colleges, high schools and elementary schools are to be able to profitably use the computer for instruction.

The organizing procedures followed by a group of chemical engineers interested in using the computer for instruction, is a model which might be usefully followed by other discipline-oriented groups of users. The CACHE Committee (Computer Aids for Chemical Engineering Education) which is composed of twenty professors of Chemical Engineering, maintains a liaison in each of the departments of Chemical Engineering in the continental USA. The group has transported much computer-based instructional material, has made two instructional packages available over the ARPANET, has published a guide to chemical engineering computer-based instructional material, and is now negotiating with industries to obtain industry developed packages like FLOWTRAN for academic use. Members of the CACHE Committee agree that these same procedures could easily be followed by instructional users in other disciplines.

Chapter 13

Instructional Computing Case Studies

Computers in Instruction

by Kenneth L. Bowles
University of California, San Diego

The University of California, San Diego, is a case study in the area of instruction and particularly in the use of batch facilities. At UCSD computer center staff have come to the conclusion that it's going to be some time before a network of the ARPANET type is going to provide the highest quality service in this area. Rather one must look for some economies of scale in a slightly different direction not disconnected from the ARPANET, but not using it initially.

Let there be no misunderstanding. Timesharing and the use of computer-aided instruction are valuable. However, there is an important element of computing which can be carried out best through the use of very fast turnaround, batch services. Some economies of scale in timesharing can be achieved either with the many small machines or through a large-scale facility like the Dartmouth timesharing service. UCSD has followed in the footsteps of a group at Case-Western Reserve University in providing for *very* fast turnaround, low cost, batch service to users. Fast turnaround in this case means under two minutes, particularly well under two minutes.

Various universities are using software like WAT-IV, WAT-V, and ITRAN in batch systems to reduce costs by running small student jobs together. UCSD has attempted to do this and presumed that the use or the growth of the computing load would be sufficient to make the batching worthwhile. In 1973 at UCSD 7,000 students in 200 courses make use of the computer predominantly for homework examples in a particular discipline. Timesharing may not be as popular at UCSD because it is not as finely tuned a timesharing system as may be available at other locations. Average cost range for timesharing is $4-$5 per hour while the typical student job can be run in a batch facility at a cost of approximately 10¢. It is possible to give the student 50 input cards and one hundred lines of output with less than two minutes turnaround by batching at a cost of approximately 10¢ per job. If a student runs 5-10 jobs per week for a particular course, costs per student, per course, per week are about $1. Furthermore, since the student gets a much greater volume of output from the batch system in two minutes than he would from a timesharing terminal, the instructional value of what he gets in this way is greater than it would be with a timesharing system in many cases. The increase in instructional value with hard copy print out is substantiated by a study made several years ago by John Skelton, University of Denver.[1] Dr. Skelton compared, under controlled conditions, the instructional value on a timesharing system to that received on a batch system. With one half hour turnaround there was a clear statistical advantage in favor of the batch system. Students were more enthusiastic at the end of the course about what they learned and they learned much more thorough methods of programming.

Conclusions like these led UCSD to pursue the implementation of batch system although it is also clear that there are some interactive systems such as APL and BASIC that fairly can't be touched in a batch context. At UCSD batch student service is no longer offered at 10¢ per job for one simple reason. With 7,000 students on the campus and five years of experience, an equilibrium has been reached.

The UCSD Computing Center is open for the students 14 hours a day in practice. If one assumes a 14 hour day (840 minutes), the use of at least two languages (ALGOL and FORTRAN), and small batches (1.3 - 1.5 jobs), overhead for software development, maintenance and accounting must be large. A charge of 10¢ per job does not recover costs but the current charge of 35¢ per job compares favorably with systems that charge as much as $1.00 per student job of the same size. We have been trying to find out through the use of network technology if economies of scale justify going into networks. It should be possible to batch and enjoy those benefits running perhaps 10,000 not 1,500 students per day, which would require working with a student body of 50,000 or 100,000 students either in a large university system like the University of California with about 110,000 students or cooperating institutions in a regional network. This could be done. In fact UCSD staff have been promoting the idea that some institution in the University of California system ought to take on provision of very fast turnaround batch computing services. A network as general as the ARPA network would be ill suited for such service primarily because of the expense of switching computers, the software support, and the large amount of processing and software necessary to make the interface from the UCSD system capable of dealing with the ARPA network.

Some of the trade-offs of using ARPANET are noted by Michael Sher in the Fall 1973 EDUCOM Bulletin in which he discusses the University of Illinois ANTS system and remote use of the UCSD facility, the 360/91 at UCLA, and the MULTICS system at MIT over the ARPANET. For sophisticated system programmers, at the University of Illinois, relative to total computing charges the cost of network transmission is 2%-3%, connection is 5%, and remote software, hardware is 7% yielding a total of perhaps 15% additional cost including the investment in a retail outlet small computer in Illinois. Networks are economical for that class of service. However, for the economies of scale gained by running small student jobs, one might well look instead at an extension of the remote job entry terminal network of a type similar to the one at UCSD. A card reader-line printer terminal, to run 200-250 cards per minute, and 200-300 lines per minute can be acquired with line and modem costs within a large state like California at an average of $1,000 per month. Including costs for paper supplies, keypunches, and so on, $2,000 per month will be adequate to run an RJE terminal capable of handling approximately 1,000 student jobs per day. Such a system might in fact have its own star network. Since many RJE terminals can operate effectively 20 days or 20,000 students jobs per month, the cost for having the terminal itself could be 10¢ per job. One could get the cost per student job in a network with a high volume down to 20¢ and, by limiting the size of the jobs allowed to run through the job stream to serve those terminals, one could get the net cost down to between 10¢ and 15¢ per job.

This type of computing service ought to be operated much as the library is in the university where the institution pays more for accounting for the work done by individual students than it does for the work. One can take advantage of the fact that the RJE terminal has a saturation limit in the number of student jobs that can be run. If one limits the size of a student job to consume no more computing resources than 15%-30%, then the average cost will be approximately 1/3 of that limit. If the overall budget pays for these jobs, twice as much computing value can be delivered to the students.

REFERENCE

1. John F. Skelton, *Time-Sharing Versus Batch Processing in Teaching Beginning Computer Programming: An Experiment*, research publications MS-R-7132, Department of Mathematics, University of Denver, 1971.

Free-Access Computing at Dartmouth

by Arthur W. Luehrmann
Dartmouth College

There is a long-standing controversy regarding the proper method of allocating computer resources to the academic community. At one extreme, exemplified in varying degrees by most universities, computer use is treated as a marketable good, and allocation is by fee-for-service, on a pay-as-you-go basis. The ultimate consumer, whether it be a research professor with grant money, an academic department with a budget, or a student with his own funds, checks his pocketbook, and then decides how much computing to buy. Open competition in a free market determines both aggregate demand for and individual allocation of computer use. At the opposite extreme, represented at Dartmouth College and at a few other universities, computing is regarded as a good that is *priceless*, in the technical sense that economists use the word. Computing is a good whose subjective worth is extremely difficult for an individual consumer to estimate in advance. Library use is an excellent example of a priceless commodity, and it is so administered at almost every educational and research institution. If the annual cost of the library had to be recovered by means of a borrowing fee, the price would be unbearable, more than ten dollars per circulation at Dartmouth. Such a policy would drive circulation down and increase the borrowing fee still further, until library use finally became the exclusive property of a tiny group of narrowly professional users who had grant money or line items in a departmental budget.

Instead, almost all institutions allocate library usage by granting free and open access to all members of the academic community. A free-access computer policy is simply an application of the "library model" to computing. .

Administrators at Dartmouth have been told for many years by those who favor a fee-for-service policy that a free-access computer policy would lead the college to financial ruin: that if Dartmouth survived, it must be a very atypical university (wealthy, small, non-technical, not research-oriented, etc.) and a poor model for others. In recent months, Kiewit Computation Center staff have been collecting and analyzing data on usage of the Dartmouth computer, in order to understand better what has actually happened after ten years experience with a free-access policy, and to answer some of the questions raised by skeptics. In sum, the data show:

- The main effect of a free-access policy is that *nearly all* members of the community use the computer.
- A *small fraction* of these people account for a very large fraction of the total usage.
- The Dartmouth "big user" community is *not different* from that at other universities and includes students as a small minority.
- If one accepts as an inevitable cost the need to supply computer service to the "big users", then the *added cost* of a free-access policy for everyone appears to be no more than a twenty to forty percent increase in aggregate demand, while the *added benefit* is a tenfold to twentyfold increase in the size of the total user community compared to the "big user" community.
- Free access does *not* mean that supported research cannot be charged for computer use.

The history of computing at Dartmouth need not be reviewed. It is enough to say that, from the very start, the primary justification for having a computer at all and for increasing its capacity and the attendant *costs* has been the belief that a knowledge of computing would add *value* to the education of students far in excess of those costs. The College issues to every student, faculty member and administrator, a plastic, wallet-sized identification card with his or her name and an I.D. number embossed on it. That number also represents each individual's personal computer account number, or "user number" in the local jargon. It is all that is needed in order to log into and use the computer although password protection is also available to each user. Possession of a user number is the right of every member of the academic community. An individual's decision to use the computer is not constrained by a concern for costs to be accounted against a personal or department dollar budget. There are no budget items for computer use at Dartmouth, although all use is accounted for and externally funded projects are charged actual dollars.

It is essential to distinguish carefully between free access and unlimited use. Free access means that any individual sitting at a terminal may dial the computer and log in without seeking either funds or permission. It does not mean that the user may execute programs that consume arbitrary amounts of computer time or file storage. As the economists say, if one forswears money then some other rationing principle must take its place, the total resource being finite.

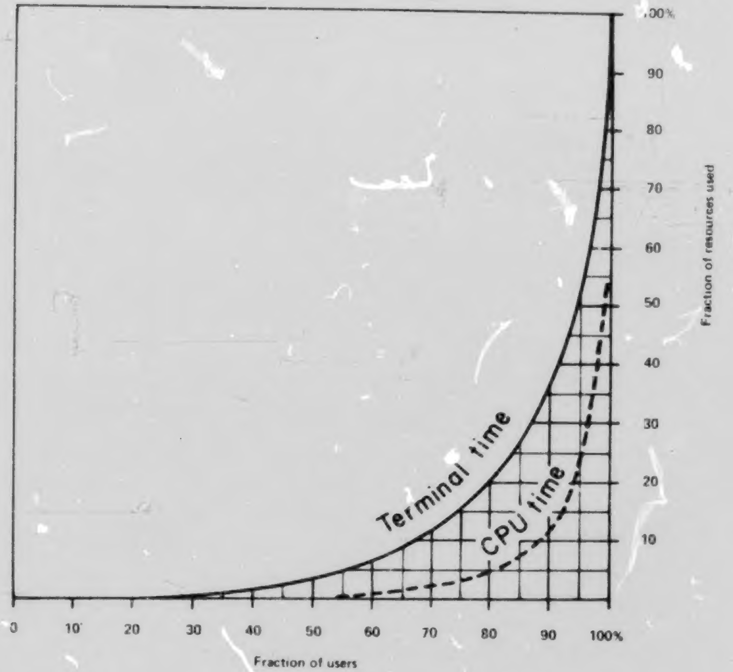
Rationing of computer resources at Dartmouth is essential and exists in two fundamental forms. The first arises from the very nature of a large time-sharing system. On a typical afternoon, when 150 people are in simultaneous contention for machine resources, even the most abusive user would find it difficult to consume more than a few percent of the total resource. Thus, time-sharing is intrinsically self-rationing in a way that batch-processing is not. The second form of rationing is explicit and requires a modest amount of administrative supervision. A feature of the Dartmouth Time-Sharing System is the ability to establish for each user number a set of specific limits under which that person must work. A typical student, for example, is limited to: 32 seconds of processor time, "CPU seconds", per job executed; 16,384 words of core-memory during execution; 20,240 bytes of long-term file storage; and no access to the card reader or punch at the computer center. A faculty member has limits also, though somewhat more generous. Such a system would not be workable if the limits were rigid, since there are times when some users need more than their current allotment. The DTSS software makes it a 30-second task to change the limits of an individual or a group, and many such requests are received and evaluated by a computer-center staff member each day.

The Dartmouth model, once more, is the library. Anyone should be able to browse through the catalog, use the reading rooms, and borrow an arm-load of books; but, not everyone should be admitted to the rare-book collection; and no one should be able to drive a truck up to the loading dock and haul off ten percent of the entire collection.

It should be understood that the Dartmouth computer has a capacity and per-capita demand similar to that at other universities of very different characters. Hence the data and analysis should not be dismissed merely because of a presumption of gross differences between systems or aggregate levels of demand. Aggregate usage data tells only part of the story and can be extremely misleading. May, 1973 usage data showed 31,499 terminal-hours and about 4,000 active user accounts. This might invite one to summarize by saying that the "representative user" uses about eight terminal-hours per month. However, such a summary would be a poor way to comprehend the actual data. In fact, the very idea of a representative user is wrong.

Each month the computer produces a summary of all activity for each user account. In any month four to five thousand accounts show some activity. Kiewit Computation Center staff have been studying the terminal time and CPU time used for each user account. After examining the usual histogram displays showing the number of users in increasing usage bins and after some experimentation, the staff hit upon a much more revealing presentation. Consider the terminal-time data, for example, as a simple list of four thousand or so numbers, one per user. Now, suppose one sorted the list so that small users came first. One might graph the result immediately, but a somewhat better picture emerges if an additional step is taken; each person's actual usage is replaced by the *cumulative usage* due to that person plus *all others who had a lower usage than he.* The result for May, 1973 data is shown as the solid line in the Figure 13.1. Note that the points along the horizontal axis represent all the four thousand users. Ascending the curve from the left, one can see at any point what percent of the total usage was due to what percent of the users. The striking

Figure 13.1 Cumulative Usage of Kiewit Computation Center, May 1973



feature, undoubtedly, is the fact that only a tiny fraction of the total usage was consumed by a large majority of the users. For example, 50 percent of the users (about 2,000 people) consumed collectively only 3 percent of the terminal time. Viewed from the other end of the spectrum, the data show that the top 5 percent of the users (about 200 people) used 50 percent of the man hours spent at terminals. The dashed curve shows a similar graph of cumulative CPU time, for which the skewing is even more evident. Here the 200 big users consumed 75 percent of the CPU time; and the 2,000 small users used only 1 percent of the CPU time.

It is fair to say that under the Dartmouth free-access policy a substantial majority of users has negligible collective impact on the total resource, even though one must assume that each user is satisfying 100 percent of his or her computing needs. Most users appear not to need very much. Obstacles erected to casual computer use would succeed in alienating several thousand people and only regain a few percent additional capacity for the serious users.

Computing Services at Chicago

by Fred H. Harris
The University of Chicago

The University of Chicago is a private, nondenominational, coeducational institution of higher learning and research which puts equal emphasis on research and teaching. It has approximately 7,500 students, of whom 2,100 are in the undergraduate college and the rest are enrolled in the four graduate divisions and seven professional schools. The University of Chicago has attracted a strong independent faculty of international distinction which numbers approximately 1,100, and there are a comparable number of research associates, lecturers, field workers, and other supporting professionals.

The Computation Center at the University of Chicago is a centralized activity with responsibilities for providing computer-related services on a cost recovery fee basis for research, instructional, and administrative purposes. Income for services rendered, whether from grants and contracts or from the University budget, is derived from discretionary monies which the responsible individual may allocate to other purposes. While the University community is in some ways a constrained marketplace, alternatives to the use of the Center's services do exist and are used. Thus the Center is expected to offer a broad range of services on a competitive basis. Increasingly, University computing services must be of reasonably comparable quality to alternatives but offered at lower marginal costs.

Cost recovery is a basic operating policy of the Center, and the desire to retain maximum support from external funding sources is a strong one. The University administration and Board of Computing Activities and Services has traditionally rejected the concept of free access. However, several years ago when external support of computing declined, the University increased its support to Computer user groups through the regular budget and continued to underwrite a deficit computer center operation with a direct subsidy. In the academic year 1972-73 sources of income to the computer center by percentage were: 21% from government and private grants and contracts; 77% from the University; and 2% from other external sources. Figure 13.2 shows sources of funding for each of three years 1970-71 through 1972-73. On the average, annual expense for the Center over the last several years has approached 4% of the University's general academic budget. Few universities support their computing centers to this extent.

Since it opened in the fall 1962 with an IBM 7090, the Computation Center has successively used an IBM 7094 coupled with a 7040, an IBM 360 Model 50, IBM 360 Model 65, and has recently installed IBM 370 Model 168. The Center traditionally has offered batch processing services with emphasis on turnaround as a function of price-related priority and resource requirements. Over the years access has been made more convenient through the addition of high-speed remote job entry stations and low-speed remote job entry facilities such as WYLBUR. A special priority was also implemented several years ago for nonsetup jobs with limited core and CPU requirements like student debugging tasks. Typically turnaround time for this type of job is extremely good. The resulting service, as viewed by the user, is illustrated by turnaround data, the relative use of high and low priority and the growth in use of remote job entry stations and online disk storage. Figures 13.3 through 13.6 illustrate the trend of these indicators from July 1970 through June 1973. To make the computer even more accessible, the Center will be implementing time-sharing services in 1974.

After considering several alternatives for future access to computing resources, the University Board of Computing Activities and Services decided in 1973 that it could best expand services over the existing base of support by purchasing a compatible larger computer system.

The marginal cost of wholesale acquisition of blocks of time from major centers off campus would be greater than the marginal cost of acquisition of the next larger system for the Center.

Finally, in the judgment of the Board it will be at least five years before computing networks of the ARPA type will be available on a viable service basis with organizational and funding problems resolved. The University's decision to amortize the 370/168 over ten years sheds additional light on the Board's estimate of when new alternatives with major improvements in cost effectiveness will be available. Our justification analysis indicated that

Figure 13.2 Computation Center
Sources of Funding

SOURCES OF FUNDING	FY 1970 - 71	FY 1971 - 72	FY 1972 - 73
Restricted Accounts (Government)	23.7%	16.5%	15.6%
Restricted Accounts (Non Government)	6.0	6.5	5.2
Subtotal	29.7%	23.0%	20.8%
University Budget	54.1%	70.7%	69.0%
Deficit	12.7	3.0	8.2
Subtotal	66.8%	73.7%	77.2%
External Accounts (Academic)	2.3%	2.2%	1.7%
External Accounts (Commercial)	1.2	1.1	.3
Subtotal	3.5%	3.3%	2.0%

Figure 13.3 Median Hours Turnaround by Month

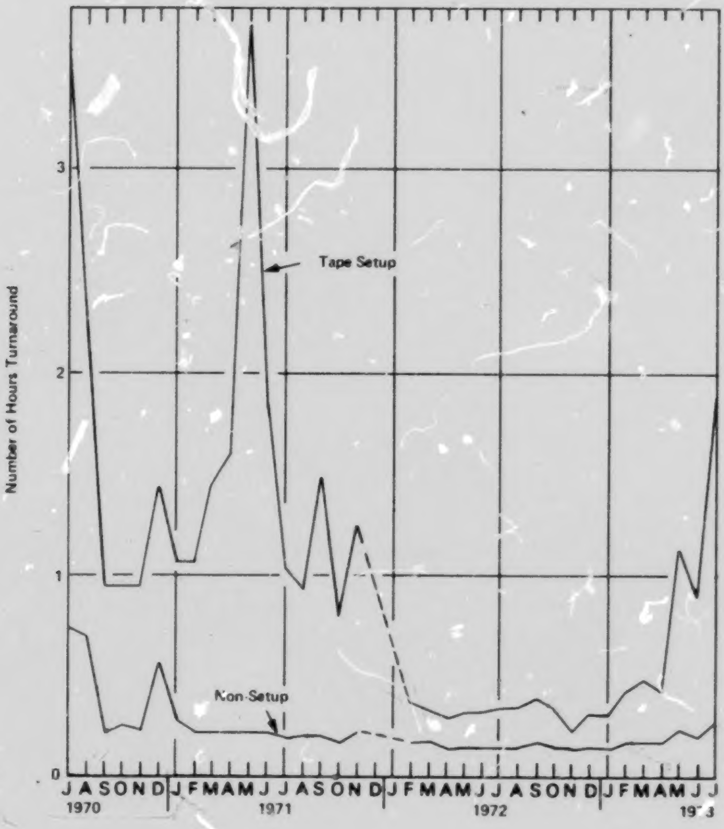


Figure 13.4 Relative Use of Low and High Priority on the 360/65 (Billable Usage Only)

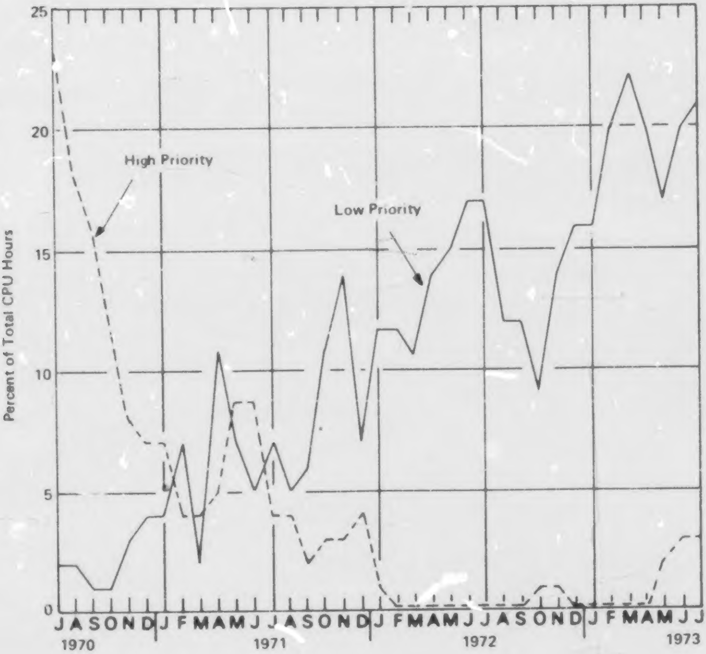


Figure 13.5 Jobs Submitted Via RJE

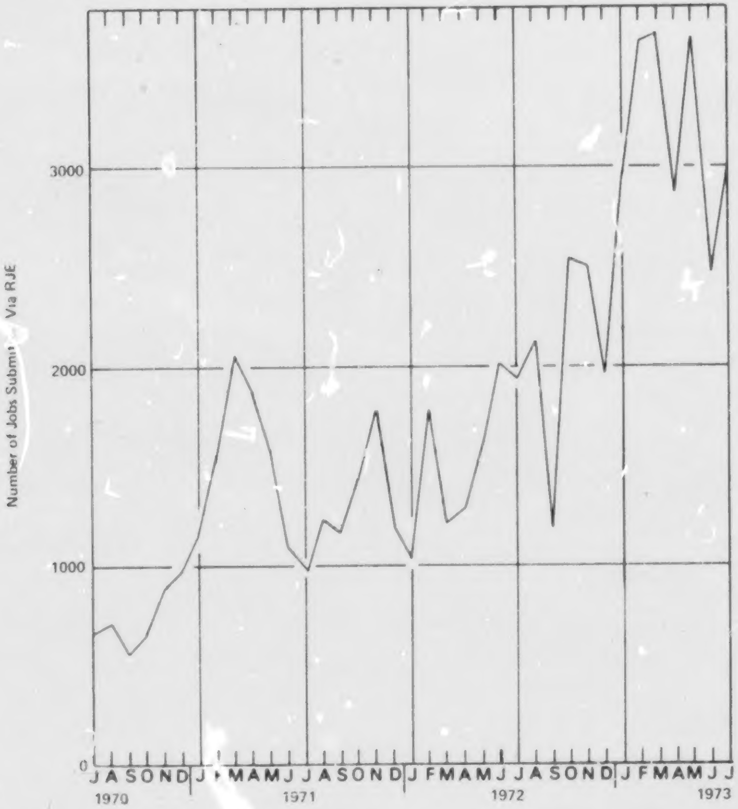
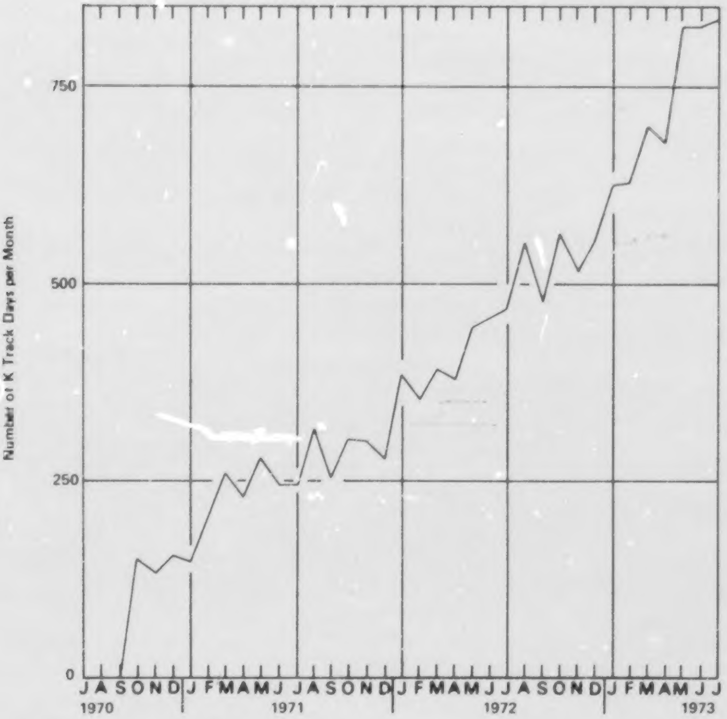


Figure 13.6 Days per Month in Which Online Disc Storage Exceeded 1000 Tracks



the 370/168 will provide the University with four- to six-fold increase in capacity for less than 10 percent increase in expense. No such price performance was available from other sources with the capacity to handle an equivalent workload. There are more than 50 laboratory minicomputers as well as several divisional computer facilities on the University of Chicago campus. Use of the minicomputers for laboratory equipment control and data acquisition will, and should, continue without conflict with the Center's services. Through a hierarchical approach to integrating and interfacing these minicomputers with the Center, (See Chapter 4) any general-purpose computation now being done on these systems should shift back to the Center.

The challenge which now faces the Computation Center is to aid in generating the demand by making the expanded capacity available in a broader, more effective way than has been done in the past. Algorithms now being considered for that purpose include: an expansion of the three-priority differential pricing system to at least a five-priority scheme; provision of volume discounts as a function of fixed commitments; and provision of matching University funds for increased external funds above some negotiated base level. While encouraging use of the Computation Center, the University administration wants to avoid wherever possible the image of undue constraints and is sensitive to the economic paradoxes of centralized functions operating with charge-back policies in decentralized environments. The uneconomic alternatives which may arise from the viewpoint of a decentralized unit are very real, and the University is turning to marginal cost analysis to resolve such conflicts. Additional budgetary support may be appropriate in individual cases.

The Board of Computing Activities and Services believes that this approach to providing major computing resources is a sound one. It effectively balances the demand for computing resources, at reasonable costs, with the other needs of the University.

Trends in Instructional Use of Computers

by Gerard P. Weeg
The University of Iowa

The Computer Center at the University of Iowa was established in 1958, with what in retrospect seem to be two fundamental mistakes: 1) the center was established essentially in support of research; and 2) the center was largely self supporting, with more than two thirds of income from non-state funds. Instructional use of the Computer Center has grown only gradually and the expected surge of instructional use of computing has plateaued at Iowa.

Several questions are under experimentation at the University of Iowa today:

- How should educational computer use be delivered?
- Should there be separate computer budgets for departments and courses?
- Will free-access lead to bankruptcy?
- Is time-sharing essential?

The University of Iowa's computing center has always been a centralized facility, co-existing with approximately 24 on-line laboratory centered computers. The central facility has been principally batch oriented, although there are 80 or 90 interactive terminals located on campus. Unfortunately, all but a handful of these terminals are private, and are not available to students.

To improve service for students a super-batch system was originated in 1972 which consists of running several special classes namely, CLASS=T, U, V, W, X, using WATFIV, PL/C, ASSEMBLER G, SPITBOL=SNOBOL, and WATBOL=COBOL. Every fifteen minutes all jobs of any of these classes are taken in from the job queue and the corresponding in-core compiler run for all jobs of each class. With the super-batch system, turn-around time for the most popular, WATFIV, is about eight minutes. With handling, the effective turn-around varies from a half-hour to an hour. In October 1973 an average load of 600 to 1,000 jobs per day was run in the WATFIV class with peaks of up to 1,500 jobs in a day. Yet even with this service, no more than 20 percent of classes at the university make use of the computer for instructional purposes. Is this low percentage due just to the fact that computing at Iowa is batch oriented? Probably not. Believing this, the Computer Center has mounted a strong educational effort to induce faculty members to see the value of computing as a supplement to instruction.

However, a study of Regional Computer Networks,[1] indicates that the mode of computing *does* affect the quantity and perhaps the quality of instructional computing. In particular, observations at Dartmouth, a university with approximately 3,500 students, showed that in peak months 1,700 students logged on the computer, and in the total year, more than 3,000 students logged on. Clearly a great number of parameters can cause such a significant involvement of students with the computer, but, equally clearly, the omnipresence of time-sharing terminals on that campus must be a contributing factor. Moreover, with a broad variety of languages available, BASIC represents 90 percent of the usage.

At Iowa, Computer Center staff and administration are rethinking the computing delivery system. Since 1970, the University of Iowa has allocated computing services through collegiate computer fund allocations which were parcelled out to departments, and further divided among individual courses and instructors. However, there has been no convenient way to parcel those funds out to the student, for whom the university exists. The system is fiscally satisfactory, but it in no way contributes to the instructional use of computing. To encourage greater use of the computer for instruction, several new tenets have been adopted:

- Instructional computing implies time-sharing conversational computing.
- Instructional computing should be *free* to the student.
- A single simple language, BASIC, will handle the bulk of the instructional need.

The interactive mode has several advantages for student use compared to batch mode. First, the student's train of thought is seldom interrupted. Second, assignments can be completed in synchrony with class scheduling. Third, the student is enthralled by his vast new control. As a result, he or she almost always does more at a session than just an assignment. Finally, the temporal proximity of results available in time-sharing seems to make insights possible which are not attained in the traditional coitus interruptions of the batch systems.

From the Study of Regional Networks, one may conclude that six teletype terminals can serve 250 students who receive one assignment per week requiring the use of the computer. If the University of Iowa with 20,000 students aims for 50 percent utilization, it follows that a minimum of 240 terminals are required to stimulate an instructional computer revolution. The university administration is prepared to plan for 1,000 terminals on campus ultimately. How can such massive computing be provided in these parlous times? First, if there is a large computer which has the potential of supporting 250 to 1,000 terminals, it would be virtually impossible to find the capital to acquire it. Moreover, after surveying the field, the Computer Center staff doubts that such a system exists. However, the minicomputer market grows faster, more capable, and cheaper as the days go by.

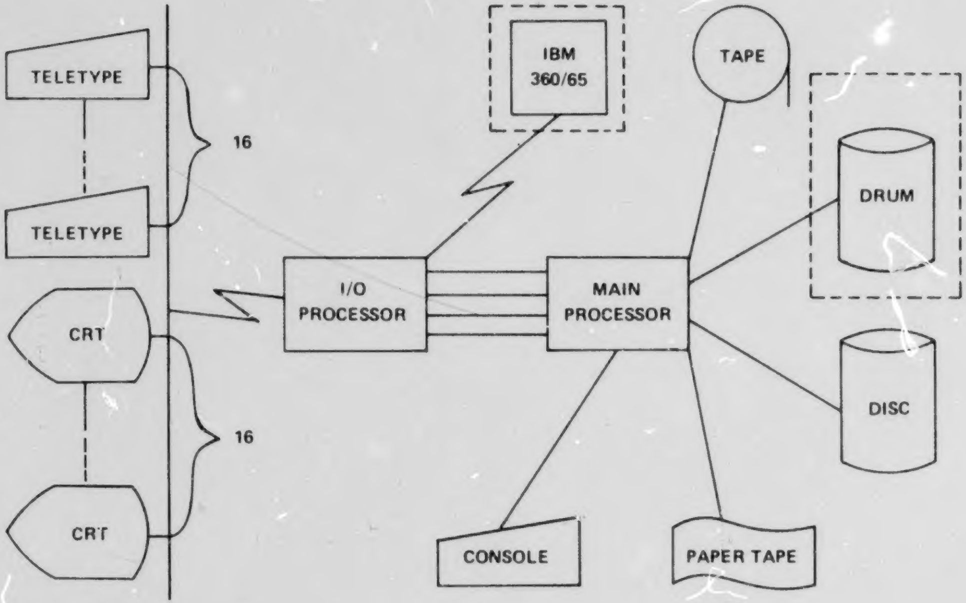
Following lots of arguments and hosts of meetings, the staff and administration at Iowa have decided to provide instructional computing through distributed computing. On campus ten, twenty, or more minicomputers, all interfaced to the central computer, the IBM 360/65 or its successor, will present conversational programming capability to University of Iowa students. The bulk of the student time-sharing load will be handled by the mini-system, and a strong effort is planned to train faculty in the use of such equipment.

At present three HP2000F minicomputers on campus should be joined by a fourth in early 1974. Thirty-two terminals in the field should soon increase to 96 in the very near future. In the spring of 1973 there were no terminals on campus.

Computer Center staff are conducting a tightly controlled experiment on the first installations. Sixteen terminals are located in the College of Business Administration and eight each are installed in the College of Education and the Department of Social Sciences. With an enrollment of about 800, the Business College is nearly computer saturated. How this college works out is crucial to continued development. If expansion plans are followed, 32 terminals will be placed in the Department of Statistics for use in consolidating two dozen introductory courses in Statistics, and 16 will go to the College of Dentistry with sixteen more assigned for general use. The cost of this kind of computing per se is around $5,000 to $6,000 per terminal period.

The minicomputers are all located in a new computer building immediately adjacent to the IBM 360/65. With hardwire connections each minicomputer can interface the 360/65 with very high transmission rates. The physical connection, as shown in Figure 13.7 is straightforward, but the software interface is a task of some significance. The connection is intended to provide file back up, RJE into the batch system, and access to the time-sharing system of the 360/65.

Figure 13.7 University of Iowa HP 2000F to IBM 360/65 System
(The drum is present on only one system)


TELETYPE
TELETYPE
16
CRT
CRT
16
IBM 360/65
I/O PROCESSOR
MAIN PROCESSOR
TAPE
DRUM
DISC
CONSOLE
PAPER TAPE

In each of the Colleges of Business Administration, Education, and in the Social Science Department deans have established locally responsible committees to: determine autonomous operating conventions; promote educational use of the terminals; spur faculty retraining; and establish computer use objectives, measurement techniques for evaluating the degree to which the objectives are met, and the time scale for meeting them. In the Computer Center, in addition to the technical team of five full-time-equivalent staff working on the software interface, a liaison team consisting of three learning specialists, a communications specialist, and two operating specialists assist faculty. With the terminals in place less than a semester, visible results have been obtained. In Social Sciences, 22 active faculty users, with four classes enrolling 325 students, plan to use the terminals in the second semester. In the College of Education, four major teaching packages have been completed, and two data banks are under construction. A large number of other projects are planned. In the College of Business Administration, 261 account numbers have been opened, 165 of which are individual undergraduate accounts. Six courses are being taught in the Fall 1973 semester, with an enrollment of about 300 students, and it appears that early in the semester a mean of 11.4 terminals out of 16 are in use regularly.

Thus a simple start at the University of Iowa is having immediate pay-offs. As progress continues, it will be reported.

REFERENCE

1. Weingarten, Nielsen, Whiteley, Weeg. *A Study of Regional Computer Networks,* (University of Iowa, Iowa City, Iowa) 1973.

Chapter 14

Transportability of Instructional Systems

Networking Challenges: The User's Viewpoint

by Thomas N. Pyke, Jr. and
Robert P. Blanc
Institute for Computer Sciences
and Technology,
National Bureau of Standards

In this paper a number of problems that impede the effective sharing of computer and information resources are identified and discussed. Taking examples from the use of present research and operational resource sharing networks, the difficulties associated with measuring and comparing performance of services provided, identifying and comparing costs to the end user, and determining the amount of effort required on the part of the user to successfully utilize a computer network are presented. Some approaches toward the solution of these problems are also discussed.

Recognizing the many benefits made possible through the use of computer networks, it is both interesting and worthwhile to identify and discuss some of the challenges that are still faced in the development and use of networks. In this paper, attention is given to computer networks in which users may access computer resources such as shared large or special purpose computer equipment and specialized software packages, as well as, shared data bases or information resources. Resource sharing networks in this general sense also provide a basis for the sharing of people as resources both in accessing and in being accessed through computer based networks.

Examples of this type of network range from the very successful centralized computer resource sharing effort at the Triangle University Computation Center in North Carolina to the large experimental general purpose resource sharing experiment sponsored by the Advanced Research Projects Agency of the Department of Defense. Many of the concerns expressed in this paper have been derived in part through the successful use of the ARPA Network at the National Bureau of Standards and through extrapolation of the use of this type of technology in a more open and widespread environment. Included in this category of networks are those large commercial computer networks that provide a terminal user access to several large computers through a terminal-oriented computer-communications network such as TYMNET and the GE network.

It is important to understand the prospects and particularly the current limitations from a user's viewpoint so that prospective users can fairly evaluate and compare resource sharing networks and network services and be aware of potential pitfalls. It is also important that investigations of solutions to the problems discussed here be promoted so that the full potential of computer networks to share expensive resources and provide equality of access to equipment, programs, people, and data can be realized.

USER SELECTION OF NETWORK SERVICE

Consider a user having a requirement for service from a resource sharing computer network. How can the user's requirements be stated clearly and even qualified where possible? What is the basis for comparing available services? Factors considered are applicable to the implementation of special purpose or in-house networks as well as to the acquisition of network based services, since these factors are oriented toward performance, cost and usability of a network.

A user may require access to an interactive system for program editing, debugging and execution: to a particular application package or to a variety of applications programs; to a particular data base or a group of data bases. It is difficult for a prospective user to clearly specify functional needs, much less consider how well services are provided through even minimal performance criteria such as response times and throughput. It is important to note that the user (presumably the person who sits at a terminal) is associated with a customer, the organization that pays the bill. The customer organization is very likely concerned not with the response time the user sees or with what programming languages a remote network based service provides, but with how much the network, together with the customer organization's staff, is better able to perform its mission. A customer organization may be concerned with potential loss of control over the resources necessary to perform its mission and with other political considerations, factors which may not be of concern to the actual user, interacting through the network with computer-based services.

The user is concerned with the service received from a remote computer system as viewed through a supporting computer-communications network. Except for crude measures such as "average response time", few meaningful criteria have been identified for indicating the performance of a computer network. Given this lack of identification and understanding, it is impossible to generate performance-oriented specifications in any quantifiable manner either for formal internal comparison of network services or for formal procurement action.

Performance measurement techniques that indicated percent utilization of a CPU or some other internal measure of performance are meaningless to a remote network user unless payment is for entire blocks of time on a remote host computer. The user should be more concerned with the amount of work performed per unit time and the cost for the entire job or for various subsets of that work. Examining performance from the user's viewpoint leads to external measures of performance and to new measurement techniques, some of which are now being investigated at the National Bureau of Standards. Techniques currently being investigated include recording the dialogue between a remote terminal user and the host computer system through the use of a new tool, a "Network Measurement Machine".

Connected at a point between the user at a terminal and the terminal interface to a network, the Network Measurement Machine can identify and time tag each character in the user/system interaction. Resulting data can then be interpreted to provide the analyst with two important kinds of measures. First it is possible to measure and comparatively evaluate system and network response time characteristics. Second it is possible to measure the demand placed on a communication network and to characterize the various kinds of workload placed on remote computer systems by terminal users. Work at NBS also includes plans for a "Terminal Environment Simulator," in which the statistical and specific results from applications of the measurement machine will be applied in simulating all or part of the terminal environment of a remote computer system accessible through a network. Testing under such controlled load conditions will result in measurement of response times and costs associated with network-based computational service on an easily repeatable basis.

Given even crude measures of performance the user's attention can be given to cost-performance considerations. Each network service has its own price structure. Although prices and price structures vary from service to service, measures such as terminal connect time, CPU seconds, core utilization, and amount of file storage are frequently the basic components of such pricing structures. Unfortunately, it is possible for such basic components to bear little direct relationship to the end service the user receives. For example, a user is charged the same for terminal connect time regardless of the loading of the remote computer system to which he or she is connected. Clearly, there will be times of day and days of the week when loading is heavier and response time poorer than other times. More connect time is required under such heavy load conditions for the user to accomplish the same sequence of functions. The user then pays more for this cost component, even though unit cost has remained the same. Likewise, a user's charge for CPU time frequently includes a portion of that CPU time required for supporting services and perhaps some overhead functions. Under heavy load conditions, where a user's program may be swapped in and out of core several times, a user may be paying for more CPU time then he would under light loading conditions. A user may pay twice the CPU charge to run a program under heavy load conditions than at other times.

Another element of cost to the user follows from what may be termed "hidden costs". There are a variety of costs not directly chargeable according to current price structures for network service. The costs for making arrangements for network service, including setting up separate accounts on two or more host computer systems in a network are frequently not obvious. The costs of interfacing a terminal, and especially a computer, to a network can be overlooked. One element of these hidden costs may be described as the effort required by the user to access, learn to use, and then use regularly, a network service. Accessibility first implies interfacing of a user's terminal to a computer network. This may be on a dial-up or hardwired basis, or may involve special equipment. The user may have to be directly concerned with character sets, character echoing, terminal speeds, and other factors. If a resource sharing network user is connecting a host computer to the network either to assist in his terminal access or to make his computer available to others connected to the network, there are additional interfacing issues. First a user must determine the requirements for the hardware interface that connects his computer to the network. In some cases a special interface is necessary, and in others the network can connect to normal I/O channels. If special purpose hardware is required, it may already be designed and can be purchased off-the-shelf, or it may be necessary to have it designed for the user's particular computer installation.

Software is of even greater concern, since some type of "network control program" is probably necessary within the user's computer to enable connection to the network. It is necessary to know if a network control program is currently available for the user's computer. If so, is it thoroughly tested, fully maintained, and is it free or must it be purchased? If it is not available, then the user ought to know how much effort is necessary to implement such a network control program, and also if there are various levels of implementation possible depending on the features desired.

Another important question is, how is such a network control program integrated within the user's computer system? If it is integrated tightly within the operating system, what happens with changes such as new releases of the operating system? What overhead is introduced into the computer system for connection to the network? Does the network control program take core or disc space? Does the network control program utilize substantial CPU time or other system resources? In some networks aminicomputer-based system can ease the burden on a user's host computer for connection to a network by minimizing the special hardware required and minimizing the new software development and software overhead required for network connection.

NETWORK DOCUMENTATION

The questions of how easy a network is to learn to use and to use on a regular basis can be explored more fully by a discussion of another major problem area, namely, documentation. Several kinds of documentation and associated consultative assistance is desirable to support a network user. General documentation should be available that clearly describes the extent and overall capability of the network or networks to which a user may require access. There must be a way for the user to determine whether appropriate resources are available on the network(s). Once one knows they are there, and is able to access them, one needs thorough documentation on available specific programs, programming language systems, and data bases.

There should be clear, well-advertised procedures for obtaining detailed network and host computer system documentation. Assistance to the user in the form of documentation and consultative service may be available by telephone, message drop within the network, or by mail. This service, including the documentation, may be supported locally, so that documentation can be brought to a user immediately by a personal representative of the network service organization. In some cases, however, it may be necessary to utilize a very remote consultant and to live with substantial delays for transmittal of documentation.

FRAGMENTED SERVICE TO NETWORK USERS

Frequently the total set of services necessary for the user to make use of a computer network is fragmented. One may have to face multiple suppliers before even making initial access to a network. For example, one may have to purchase or rent a terminal from one supplier, and then select from alternative ways of connecting the terminal to one or more networks. Modems and various protective devices may be required for connection to common carrier communications equipment. One may require a dial-out line connected to the public switched network, or it may be more advantageous, to connect on a leased circuit basis. Once access is established to one or more networks, different phone numbers must be remembered along with completely different procedures for using different network services. Even with a single network different host computers may have completely different log in conventions. The problem only begins at log in, however, since the operating system control languages would very likely be quite different. Application programs, language processors, or other packages may also have quite different conventions from a user's viewpoint.

Multi-network access is of considerable interest, since a typical user may have need for a wide range of application programs, both computationally based and information based. The wide range of different access procedures at all levels makes the task of using applications-oriented packages, on even one network, difficult. Efforts are needed to standardize access procedures within single networks and across a number of networks where possible. Where this is not possible, assistance can be provided the user through a "Network Access Machine," which can help the user by some combination of prompting or even automatic execution of access procedures on behalf of the user.

A Network Access Machine presently being developed at NBS acts as the connection point for a user at a terminal, and in turn, establishes the access path to the desired network, computer system on that network, and application package. This is accomplished through the execution of a user defined network command language, which acts to expand user commands into a command sequence, executable on a particular network and computer system. Conditional expansions allow for the use of the same commands on different networks and computer systems, while system libraries of commonly used command sequences combined with user profiles, add to the utility of the Network Access Machine concept.

SUMMARY

Problems in selecting network service can be removed if appropriate performance criteria and measures can be developed and if price structures can be changed to better correspond to actual services received by the user. Efforts are required to assist users in identifying hidden costs and hidden tasks required by them to access, to become familiar with, and to use computer networks. At the National Bureau of Standards, under NSF sponsorship, computer communications networking technology has been extensively reviewed, and the pros and cons of selected approaches have been identified and comparatively evaluated. In addition, there has been a substantial effort to identify cost factors in the use of computer communication networks. These cost factors have been combined with hypothetical user traffic demands, and selected terminal and host computer configurations, to arrive at a set of estimated cost figures for using computer communication networks.

Better network documentation is also required, but should be based on a better understanding of the purposes of documentation. A user may become an expert on one system, but may be an infrequent user of several other services concurrently. One thus needs documentation easily and quickly available, together with the necessary consultative service, in order to avoid the hidden costs associated with time wasted in getting familiar with a particular system.

The problem of fragmented service has been partly overcome by some network organizations through providing, on a rental basis, a terminal with acoustic coupler to a user. However, this may not help a user who requires access to more than one network. Terminals required for different networks may even have different character codes. Further work is necessary to provide assistance to users in the selection of terminals, and in the prospective use of Network Access Machines and other mechanisms, for making computer networks more hospitable to their users, and for making users more comfortable with networks.

CAI Programs for Multi-University Use

by Keith A. Hall
The Pennsylvania State University

Since 1970 computer-assisted instruction has been used at The Pennsylvania State University as a means of providing inservice training for teachers widely disbursed throughout the state. Developmental efforts have been focused on stand-alone courses to provide the complete program of instruction equivalent to certain college courses. The stand-alone concept was adopted for two reasons: 1) to meet the needs of the widely disbursed audience without having a professor on site; and 2) to provide CAI as a direct replacement in function as well as cost rather than being an add-on cost to existing instructional program.

The early work of the CAI Laboratory was done using typewriter terminals to which random access slides and random access audio tapes were eventually added. Since 1967 when the first production line IBM 1500 system was installed at Penn State, that configuration has been used and, in October of 1970 a program of mobile computer-assisted instruction (CAI) was inaugerated in the inservice continuing education program for teachers.[1]

CURRICULUM

The graduate level courses are designed to assist teachers in understanding, identifying, and remediating the problems of handicapped children. The courses all bear the generic title CARE of Computer Assisted Renewal Education. Developed by a team of faculty members under the leadership of Professors G. Phillip Cartwright and Carol A. Cartwright, the major objective of the CARE series is to teach inservice teachers, aides, and other educators how to work effectively with handicapped children in typical classrooms. Students interact with the instructional program at computer student stations. In addition to the program of instruction stored in the computer, each student uses a textbook, a handbook, and a set of materials for testing young children. When the teacher-student completes this course, he or she takes a 75-item final examination generated by the computer from a pool of more than 300 test questions covering the objectives of the course. A complete record of each participant's performance, not only on the test but on the course as well, is recorded on magnetic computer tape for summarization, marking and course improvement. Teacher-students are awarded marks by the faculty member in charge of the course, and they receive credit appropriate to the amount of curriculum included in the program.

FACILITIES

To implement the CARE program, a custom-built expandable van was fitted with a small stand-alone computer and 16 student stations (the IBM 1500 Instructional System). Each student station is equipped with a small cathode-ray tube (CRT) on which alphameric information plus a wide variety of graphics, including animated illustrations can be displayed. For response students use a typewriter-like keyboard with upper and lower case characters plus a variety of special characters, and a light-sensitive pen. In addition to the CRT, each student station has a rear-screen image projector which can display color photographic images from a 1,000-frame, 16mm film with each frame randomly accessible by the computer at a search rate of 40 frames per second. The third display component is an individual audio play/record device with randomly accessed prerecorded messages on 1/4-inch audiotape.

IMPLEMENTATION

On a prearranged schedule, the mobile CAI laboratory is moved to a school in a rural community and connected to electric, telephone and water services. Over the next seven weeks, in late afternoon and evening hours, elementary teachers and their supervisors schedule themselves for one- to three-hour sessions at computer student stations on flexible and irregular schedules to fit into the demands of their personal lives. During a seven-week period the laboratory will accommodate from 125 to 150 learners who enroll for a typical three-credit college course. The students, of course, put in considerable time in home study of the textbook and the 400-page handbook which accompany the course. Currently the existing curriculum, student stations, and relocation every seven weeks meet the needs of a target audience in a 25-mile radius.

The field staff for the program consists of a manager who travels with the mobile laboratory, plus a computer operator and two student proctors hired in the local community to help students with scheduling and student station operations. The faculty member in charge of the course is available from time to time to talk with students in person and can always be reached by telephone.

CURRICULAR COMPATIBILITY

CARE programs have been offered for credit by nine different institutions in addition to Penn State. These institutions and the enrollment at each institution for the CARE courses are shown in Figure 14.1. The extent to which the curriculum has been accepted by other institutions is a direct result of the efforts of the development team. During the development of the courses the curriculum was continuously reviewed by faculty members from other institutions, not necessarily those where the credit was eventually offered for the program, to insure a broad perspective of the curriculum concepts and to insure that the course was abreast of the current trends and patterns throughout the United States.

SYSTEMS COMPATIBILITY OR CONVERTIBILITY

CARE course materials are compatible with several systems primarily because student station devices are functionally equivalent on several systems. Only programming modifications and clerical support are required to make the conversion where two systems are functionally equivalent. However, if the functional characteristics like the graphics CRT Audio Unit, or photographic image display of the student station are drastically dissimilar, educational and instructional decisions will be required to determine new methods and techniques for restructuring the curriculum.

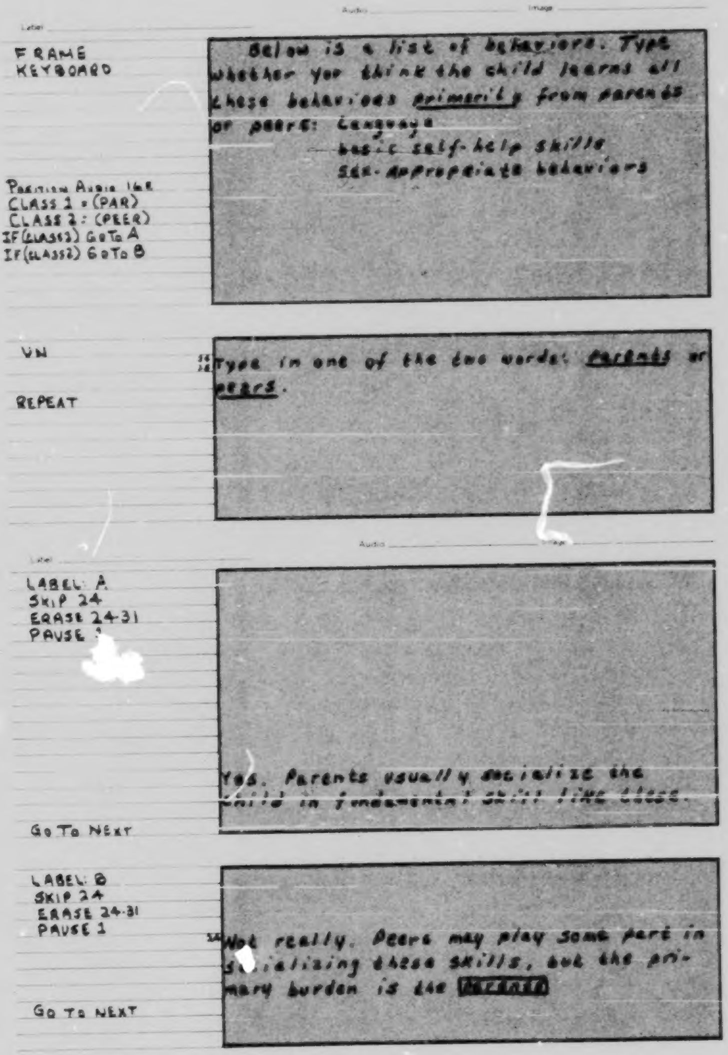
Part of the effort at Penn State has been to develop techniques which will provide a computer-system free and computer-language free documentation of the curriculum. One application program TACL (Teaching And Coursewriting Language) not only provides documentation as a by-product but eliminates the need for coding curriculum material in any computer language. TACL enables authors to develop curriculum content without knowing any computer language at all. After having been written on specified forms, material is input in the computer system using an on-line CRT, compiled by the operating system, and then made available to the

Figure 14.1 Student Enrollment Summary for CARE Courses

Institution*	Registrants	Credit Hours
Penn State University		
On-Campus	1230	3648
Off-Campus	1688	5064
University of Houston	210	630
Georgia State University	81	243
Southern Illinois University, Edwardsville	126	378
Northern Illinois University, DeKalb	180	540
Indiana University	99	297
Total	3614	10,800

*University of Texas, Austin, SUNY, Stony Brook, University of Alberta, Edmonton, and Montgomery County Public Schools, Maryland, have implemented the CARE courses on their systems. Therefore enrollment data is not readily available.

Figure 14.2 Sample of an Author's Writing


Label
Audio
Image

FRAME
KEYBOARD

Below is a list of behaviors. Type whether you think the child learns all these behaviors *primarily* from parents or peers: Language
basic self-help skills
sex-appropriate behaviors

Position Audio 16R
CLASS 1 = (PAR)
CLASS 2 : (PEER)
IF (CLASS1) GoTo A
IF (CLASS2) GoTo B

VN

Type in one of the two words: *parents* or *peers*.

REPEAT

LABEL: A
SKIP 24
ERASE 24-31
PAUSE 1

Yes. Parents usually socialize the child in fundamental skills like these.

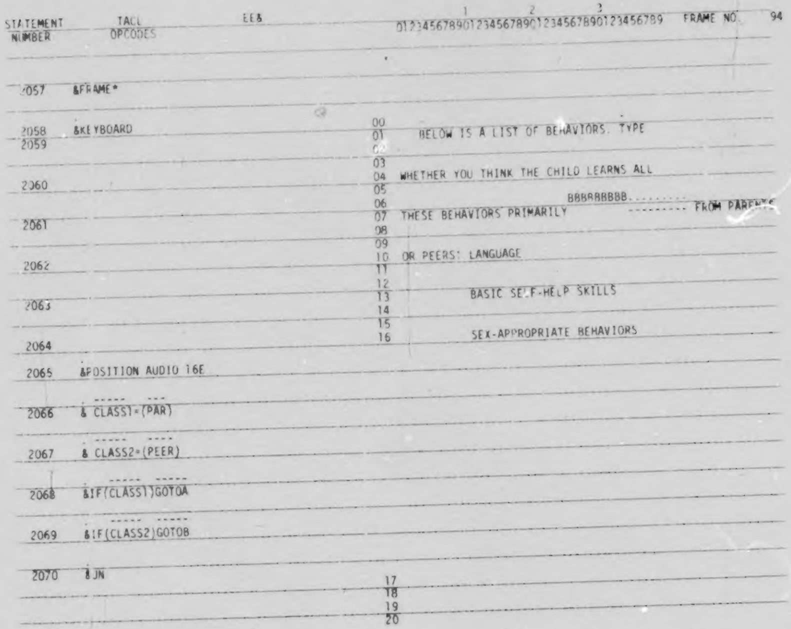
GoTo NEXT

LABEL: B
SKIP 24
ERASE 24-31
PAUSE 1

Not really. Peers may play some part in socializing these skills, but the primary burden is the parents.

Go To NEXT

Figure 14.3 Coursewriter II Program Generated by TACL



```
                                                                 1         2         3
STATEMENT       TACL                EEB                 0123456789012345678901234567890123456789  FRAME NO.   94
NUMBER          OPCODES

2057    &FRAME*

2058    &KEYBOARD                               00
2059                                            01     BELOW IS A LIST OF BEHAVIORS. TYPE
                                                02
                                                03
2060                                            04 WHETHER YOU THINK THE CHILD LEARNS ALL
                                                05
                                                06                          BBBBBBBBB.........
2061                                            07 THESE BEHAVIORS PRIMARILY          ---------  FROM PARENTS
                                                08
                                                09
2062                                            10 OR PEERS: LANGUAGE
                                                11
                                                12
2063                                            13          BASIC SELF-HELP SKILLS
                                                14
                                                15
2064                                            16          SEX-APPROPRIATE BEHAVIORS

2065    &POSITION AUDIO 16E

          -----   ---
2066    & CLASS1=(PAR)

          -----   ----
2067    & CLASS2=(PEER)

            -----  -----
2068    &IF(CLASS1)GOTOA

            -----  -----
2069    &IF(CLASS2)GOTOB

2070    &JN
                                                17
                                                18
                                                19
                                                20
```

Figure 14.3 Continued

```
                                    21
                                    22
                                    23

2071   &SKIP 24
                                    24  -                                          BBBBBBB.......
2072                                25  'YPE IN ONE OF THE TWO WORDS. PARENTS             ------- OR
                                    26       BBBBB.....
2073                                27  PEERS      -----.

2074   &REPEAT

2075   &LABEL  A
                                    28
                                    29
                                    30
                                    31
****************************************************************************************
                                                         #

2076   &SKIP 24

2077   &ERASE24- 31

2078   &PAUSE01
                                    24  -    -
2079                                25  YES. PARENTS USUALLY SOCIALIZE THE
                                    26
2080                                27  CHILD IN FUNDAMENTAL SKILLS LIKE THESE.
              ----
2081   &GOTO NEXT

2082   &LABEL  B
                                    28
                                    29
                                    30
                                    31
****************************************************************************************
```

Figure 14.3 Continued

```
 2083    &SKIP 24

 2084    &ERASE24- 31

 2085    &PAUSE01                                 24  -          -
 2086                                             25  NOT REALLY. PEERS MAY PLAY SOME PART IN
                                                  26
 2087                                             27  SOCIALIZING THESE SKILLS, BUT THE PRI-
                                                  28                 +++++++
 2088                                             29  MARY BURDEN IS THE PARENTS.
               ----
 2089    &GOTO NEXT

STATEMENT      TICL           EE&                           1         2         3
 NUMBER        OPCODES                            0123456789012345678901234567890123456789   FRAME NO.   95



 2090    &FRAME

 2091    &LIGHT PEN

 2092    &PLAY AUDIO 16E
                                                  00
                                                  01


 2093    &SKIP 02                                 02       ----
 2094                                             03       PEERS TEACH EACH OTHER.

 2095    &PAUSE03

 2096    &RESUME AUDIO
                                                  04
```

Figure 14.4 TACL Source Listing (Documentation)

```
321C  CB92
3211       1     DTI 30,36+/2,30+/4,36+/+S+P*E
3212       2     EPI 30,38+/2,30+/1,39+/+/1+/*E
3213       3     AA  (8+/)ZZ*E
3214       4     PR  +E
3215       5     DE  0+/32*E
3216       6     DTI 30,29+/2,30+/6,29+/CB92     *E
3217       7     FPO *E
3218       8     DT  4,0+/2,4+/40,0+/   (H)OW DOES SOCIALIZATION BRING ABOUT *E
3219       9     DT  7,0+/2,7+/40,0+/THIS CHANGE IN THE CHILD'S AWARENESS OF *E
3220      10     DT  10,0+/2,10+/40,0+/OTHER POINTS OF VIEW(/ *E
3221      11     PA  60*E
3222      12     FN  PSREMV+/  +/C(=)01,004L+,C2+,T(=)492L+/C(=)01,005L+,04+,T(=)541L*E
3223      13     AUP CB15*E5864,1+/234*E
3224      14     DT  18,0+/2,18+/40,0+/   (T)HE ABILITY TO TAKE THE OTHER PER- *E
3225      15     DT  21,0+/2,21+/40,0+/SON'S POINT OF VIEW ALLOWS HUMAN BEINGS *E
3226      16     DT  24,0+/2,24+/40,0+/TO LIVE TOGETHER, WHICH IS THE ULTIMATE *E
3227      17     DT  27,0+/2,27+/40,0+/GOAL OF SOCIALIZATION.*E
3228  CB93
3229       1     DTI 30,36+/2,30+/4,36+/+S+P*E
3230       2     EPI 30,38+/2,30+/1,39+/+/1+/*E
3231       3     AA  (8+/)ZZ*E
3232       4     PR  *E
3233       5     DE  0+/32*E
3234       6     DTI 30,29+/2,30+/6,29+/CB93     *E
3235       7     DT  0,0+/2,0+/40,0+/   (O)KAY, SO SOCIETIES SET UP RULES*E
3236       8     DT  3,0+/2,3+/40,0+/FOR ACCEPTABLE BEHAVIOR AND CHILDREN*E
3237       9     DT  6,0+/2,6+/40,0+/LEARN THE PROPER BEHAVIOR FROM THE*E
3238      10     DT  9,0+/2,9+/40,0+/PEOPLE AROUND THEM. (B)UT WHO SPECIFI-*E
3239      11     DT  12,0+/2,12+/40,0+/CALLY IS RESPONSIBLE FOR SOCIALIZING*E
3240      12     DT  15,0+/2,15+/40,0+/THE CHILD(/*E
3241      13     DT  18,0+/2,18+/40,0+/   (I)N OUR SOCIETY THERE ARE TWO GROUPS*E
3242      14     DT  21,0+/2,21+/40,0+/OF PEOPLE WHO ACCOMPLISH MOST OF A*E
3243      15     DT  24,0+/2,24+/40,0+/CHILD'S SOCIALIZATION(& +P+A+R+E+N+T+S )AND +P+E+E+R+S*E
3244      16     DT  27,0+/2,27+/40,0+/(9)OTHER CHILDREN OF ADULT THE SAME AGE(0). *E
3245  CB94
3246       1     DTI 30,36+/2,30+/4,36+/+S+P*E
3247       2     EPI 30,38+/2.30+/1,39+/+/1+/*E
3248       3     AA  (8+/)ZZ*E
3249       4     PRR *E
3250       5     BR  ON+/S31+/1*E
3251       6     DE  0+/32*E
3252       7     DTI 30,29+/2,30+/6,29+/CB94     *E
3253       8     DT  0,0+/2,0+/40,0+/   (B)ELOW IS A LIST OF BEHAVIORS. (T)YPE *E
```

Figure 14.4 Continued

```
CARE2-4    03/27/73

3254     9      DT  3,0+/2,3+/40,0+/WHETHER YOU THINK THE CHILD LEARNS ALL*E
3255    10      DT  6,0+/2,6+/40,0+/THESE BEHAVIORS PRIMARILY*B*B*B*B*B*B*B(---------- )FROM PARENTS*E
3256    11      DT  9,0+/2,9+/40,0+/OR PEERS(& )LANGUAGE*E
3257    12      DT  12,0+/2,12+/40,0+/          BASIC SELF-HELP SKILLS*E
3258    13      DT  15,0+/2,15+/40,0+/          SEX-APPROPRIATE BEHAVIORS *E
3259    14      AU  CB16*E5979*E
3260    15      EP  17,0+/2,17+/40,0+/+/40+/ CB94     *E
3261    16      FN  ED+/+/,D+/+/ +/,+/.+/(6+/)-+/(&+/)&+/(/+/)/+/*R+/*I+/*B+/*C*E
3262    17      AA  (8+/)ZZ*E
3263    18      LD  PAR +/B1*E
3264    19      FN2 COD+/1,1*E
3265    20      FN2 SQ+/AA+/(99)1(CO*E
3266    21      BR  CB94A*E
3267    22      AA  (8+/)ZZ*E
3268    23      LD  PEER  +B1*E
3269    24      FN2 COD+/1,1*E
3270    25      FN2 SQ+/AB+/(99)1(OO*E
3271    26      BR  CB94B*E
3272    27      UN  UA*E
3273    28      DT  24,0+/2,24+/40,0+/(T)YPE IN ONE OF THE TWO WORDS(& )PARENTS*B*B*B*B*B*B*B(------- )CR*E
3274    29      DT  26,0+/2,26+/40,0+/PEERS*B*B*B*B*B*B*B(-----).*E
3275    30      BR  RE*E
3276    31      EA  *E
3277  CB94A
3278     1      DE  24+/6*E
3279     2      DTI 30,0+/2,30+/28,0+/*E
3280     3      PA  10*E
3281     4      DT  24,0+/2,24+/40,0+/(Y)ES. (P)ARENTS USUALLY SOCIALIZE THE*E
3282     5      DT  26,0+/2,26+/40,0+/CHILD IN FUNDAMENTAL SKILLS LIKE THESE. *E
3283     6      BR  CB95*E
3284  CB94B
3285     1      DE  24+/6*E
3286     2      DTI 30,0+/2,30+/28,0+/*E
3287     3      PA  10*E
3288     4      DT  24,0+/2,24+/40,0+/(N)OT REALLY. (P)EERS MAY PLAY SOME PART IN *E
3289     5      DT  26,0+/2,26+/40,0+/SOCIALIZING THESE SKILLS, BUT THE PRI-*E
3290     6      DT  28,0+/2,28+/40,0+/MARY BURDEN IS THE +P+A+R+E+N+T+S. *E
3291     7      BR  CB95*E
3292  CB95
3293     1      DTI 30,36+/2,30+/4,36+/+S+P*E
3294     2      EPI 30,38+/2,30+/1,39+/+/1+/*E
3295     3      AA  (8+/)ZZ*E
```

author and students for testing and evaluation. Samples of an author's writing is shown in Figure 14.2. The Coursewriter II Program generated by TACL is illustrated in Figure 14.3, and the TACL source listing appears in Figure 14.4. A natural extension of TACL would be modifications to produce Coursewriter II (IBM 360/370 Systems), Tutor (PLATO System) and other CAI languages as they become available.

At the current stage of development in CAI, attention should be focused on curricula compatibility among institutions and machine readabie documentation to facilitate conversion to newer computer systems. Potential for conversion is of greater value and importance than efforts to standardize procedures, languages, and facilities as some people often urge. CAI is too young and creative for standardizing and too important and compelling to ignore futures.

REFERENCES

1. The operation of the mobil system is currently supported by the University, the Penn State Foundation, and a grant from the National Center for the Improvement of Educational Systems of the United States Office of Education. The development of the CARE courses for computer presentation is funded by a grant from the Bureau of Education for the Handicapped, United States Office of Education. Alternative strategies and a cost benefit comparison for providing inservice instruction for teachers are described in Hall, K. A. and H. E. Mitzel, "CARE: Computer Assisted Renewal Education – An Opportunity in Pennsylvania," *Audio Visual Instruction*, Vol. 18, No. 1, January 1973, pp. 35-38.

2. Cartwright, G. P. and C. A. Cartwright, "A Computer-Assisted Instruction Course in the Early Identification of Handicapped Children," *Journal of Teacher Education*, Vol. 24, No. 2, Summer 1973, pp. 128-134.

3. Countermine, T. A. *The Development and Evaluation of a Teaching and Coursewriting Language (TACL)*, Report No. R-57, (The Computer Assisted Instruction Laboratory, The Pennsylvania State University, University Park, Pennsylvania), August 1973.

Scientific and Technical Information Transfer for Education (STITE)*

by Pranas Zunde
Georgia Institute of Technology

Research and development in the field of science information during the past decade has resulted in the establishment of large banks of descriptive information and bibliographic data. Stored on digital and analog media, in science and technical information centers, these collections, along with mechanisms for their organization, search and dissemination, constitute a wide network of science information systems.

While the utilization of these science information systems has, in the past, been primarily in research, it is desirable that the use of such valuable resources be extended to other areas of intellectual endeavor. Science education, because of its inherent function of transferring information from an external source into the human mind, seems to be a natural subject for the extension of the use of these centers and systems. The recent development of technology-aided learning systems which permit learners to interact with organized learning materials stored in an inanimate manipulable device, or memory, strengthens the possibility of an increased utilization of science information for the purposes of instruction and learning.

---

*STITE is an ongoing research project at the School of Information and Computer Science, Georgia Institute of Technology, Atlanta, Georgia, sponsored by the National Science Foundation Grant No. GN-36114. Project Director is Dr. Pranas Zunde.

The objective of the STITE project is to study, design, and experimentally evaluate man-machine mechanisms for enhancing the transfer of science information from its present repositories into science learning systems. Within this general objective are the following specific and related goals:

- To describe operationally the human process of transforming science information system outputs for the purpose of integrating them into learning systems.
- To investigate comparatively the design and operating characteristics of science information systems and science learning systems, particularly from the viewpoint of requirements for transferring information between them via a man-machine interface.
- To implement an experimental design of a limited transfer mechanism from appropriate existing science information systems into a learning system and to evaluate some aspects of that mechanism.

Initial research began with an analysis of existing science and technology centers to determine the subject areas covered and the kinds of services provided to users. Learning information systems were examined from the standpoint of the processes involved in developing learning materials. On the basis of findings in these areas, some tentative conclusions were made regarding the possible functions of the demonstration model, as well as its design and operation. Concurrent with this investigation was a survey of relevant research literature to identify and evaluate materials significant in the achievement of the goals of the STITE project.

For purposes of the demonstration, graph theory was chosen as the experimental subject area. Further research has produced a list of seventeen specific kinds of tasks that might be useful for science educators which might be potential outputs of the STITE system. This list includes such functions as preparation of a course outline, presentation of illustrative examples, and presentation of a set of relevant questions on a particular topic. In relation to these possible tasks, some tentative specifications for the internal design of the system have been determined.

Further research will concentrate on specifications of design and on the preparation of data bases and programs to achieve STITE's goal of transferring stored science information into learning systems.

Chemical Engineering Instructional Packages

by Warren D. Seider
University of Pennsylvania

Since its inception in 1969, the CACHE Committee has sought a mechanism for distributing large-scale programs for computer-aided analysis and design, estimation of physical properties, and dynamic simulation and control studies, among others. While most of these programs have been developed in FORTRAN to allow for machine interchangeability, program installation and maintenance has been the major obstacle in the path of widespread usage. Most departments of chemical engineering are not equipped to maintain a large program library. Often faculty lack the necessary experience to install the programs.

The CACHE Program Distribution Task Force has been closely following the development of communication networks that link computers together. Such networks are necessary to enable chemical engineering educators and students at remote locations across the country to use any computer program installed on any of the network computers. This is especially important for widespread usage of large-scale programs, which have been developed at a single university, to be run on a specific computer, where they can be maintained by the authors.

Recently, the CACHE Large-Scale Systems Task Force completed a document entitled "CACHE Guidelines for Large-Scale Computer Programs" that describes the desirable features of large-scale programs for student use. As such, it provides guidelines for the CACHE editor who judges whether or not a large-scale program is to be endorsed by the

CACHE Committee. The guidelines are based upon the assumption that large-scale programs are maintained on a network computer inexpensively accessible to students at many universities.

At the EDUCOM Fall Conference the CACHE Program Distribution and Task Force reviewed a draft of a new document entitled "CACHE Guidelines for Computer Networks." Following a review of the computer networks available by Professor Warren Seider the workshop discussed the guidelines first as they apply to network computers and second as they apply to network communications.

NETWORK COMPUTER GUIDELINES

A network computer should provide services that cannot be furnished by general-purpose computers that compile and execute relatively small-scale programs written in FORTRAN, BASIC, COBOL, and APL, among other languages. Network computers, to justify the added costs of communications, should provide services that complement these. For example, a network computer should enable execution of large-scale programs.

Several important features for a network computer are listed below. This list is open-ended and is presented in random order; it offers a selection of features that altogether are difficult to find in most network computers.

- The network computer should have a simple sign-on procedure, with protocols that require little typing and are easy to follow.
- A reliable accounting system is necessary to keep careful records of computing done by each student and to limit the amount of computer resources available to each student. The accounting system should not allow a student to over-expand the funds in his or her account and should provide user passwords to enable a student to protect an account from misuse by other students. It should provide each school a summary of all charges at least bi-weekly and should provide an instantaneous review of the balance in each student account.
- In addition to remote batch services, a remote job entry system is necessary for communication with typewriter terminals, although a time-sharing system is preferred. The system should enable each student to enter programs and data into secondary storage and conveniently edit the information. It should provide for each addition and deletion of records and merging and disbanding of files. A string editor is desirable but not mandatory.
- Ocassionally large-scale systems display many lines of results. To reduce communication costs, especially when typewriter terminals are used, the host computer should use a line printer to print several

pages of results that are mailed to individual users.

- The host computer should, at least, have a seven and nine track tape drive and tape labeling facilities for installation and removal of programs and data. A card reader is essential. A card punch, digital plotter, and facilities for video display terminals are desirable, but not essential.
- Many large-scale systems contain subprograms that call upon themselves recursively. This is the case in information systems where a request for information initiates several other requests to be satisfied before the initial request is satisfied. When all requests use the same FORTRAN subprogram, with different parameter values, it is desirable for the program to be recursive. In these cases, a FORTRAN compiler that permits recursive subprograms is desirable.
- Systems programming personnel should be available to assist authors with installation and maintenance of large-scale programs. These persons should also be prepared to discuss problems concerning the host computers operating system and compilers with a responsible person at each of the participating universities and should provide notification of system changes before they are made.
- Literature describing the host computer operating system and compilers should be clearly stated and free of errors. This is especially important since communications with network computer personnel are usually limited to telephone conversation and written correspondence. Literature updates should be circulated promptly.
- Computation costs and connect charges should be comparable to local university computer costs. No surcharge should be paid for student usage of software developed at universities with internal or external funds. Wherever possible, a surcharge for industrial software should be waived in return for the publicity and training of prospective users.
- It should have a history of reliable performance, including a high mean time between system failures and short delay to recover from failures. Adequate back-up for on-line disk storage should be provided.

COMMUNICATION NETWORK GUIDELINES

Low cost communication networks that can be accessed inexpensively from many college campuses are required to provide a sufficiently large user base to warrant the installation and maintenance of large-scale system programs. Several factors for evaluation of communication networks are listed below:

- The network should offer 300 bps communications for video display terminals and typewriter terminals.

- The network should offer 2,000 bps communications for card reader-line printer terminals.
- The cost of communications should be considerably cheaper than standard telephone costs.
- Universities should be able to gain access to a communications network without tying in a computer. Computer-computer communications are not important for most chemical engineering classwork computations. Terminal-computer communications are important.
- The communications network should be a local phone call away from at least 30 chemical engineering departments. Its protocols for log on, log off, and communication with the host computer should be very simple and require little typing.

ADMINISTRATION OF GUIDELINES

The CACHE editor for large-scale systems will judge the suitability of a communications network for installation of a given large-scale system. He will be advised by the CACHE Program Distribution Task Force. In October 1973 several large-scale systems are ready for installation on a network computer. Fortunately, promising networks are becoming available. It is likely that during the next calendar year many chemical engineering departments will have access to large-scale computer programs and data bases on network computers.

Chapter 15

Software Compatibility in the Social Sciences

Performance Differences in Social Science Statistical Software

by William D. Slysz
University of Connecticut

One primary consideration involved in the selection and use of general purpose computer software is operating performance. This report describes the results of a series of tests applied to a set of software "packages" widely used in the social sciences, and elsewhere.[1] The packages are: BMD[2], University of California; DATA-TEXT[3], Harvard University; OSIRIS[4], Inter-University Consortium for Political Research, University of Michigan; SPSS[5], National Opinion Research Center, University of Chicago; and TSAR[6], Duke University. The most current releases of each package, January 1973, were used in the testing.

TEST DATA, COMPUTER CONFIGURATION, and COST ALGORITHM

Test data used was the SRC 1968 American Election Study, containing 533 variables for each of 1673 logical records. Logical record length was 1020 bytes, and a blocking factor of 3 was used, with the data (and packages) stored on IBM 2314 disks.

Runs were made on an IBM 360/65 computer which used IBM's OS/MVT operating system release 21.0, and HASP, version 3.1. All tests were run with other jobs in the job stream.

The costs shown in the results were derived by the University Computer Center, University of Connecticut cost algorithm:

Resource	Charge
• Central processor memory less than 145K bytes	$ .05/CPU second
• Central processor memory 145K through 160K bytes	.075/CPU second
• Central processor memory greater than 160K bytes	.10/CPU second
• Executing channel programs (EXCP's)	2.00/1000 EXCP's

(1). Eight commands to the printer, such as printing a line, spacing, etc. are evaluated as one EXCP.

(2). No charge is made for reading job or software system control cards.

RESULTS

The following graphs (Figure 15.1 through 15.6) describe the relative performances measured in terms of execution cost. There is a separate graph evaluating the descriptive statistics, bivariate tables, Pearson correlation, stepwise regression, factor analysis, and one-way analysis of variance procedures. The x axis designates the number of variables (or experiments for ANOVA) involved in a given test run. The y axis designates the ratio formed according to the rule:

$$\text{\$ Cost Ratio} = \frac{\text{\$ performance of package being tested}}{\text{\$ performance of package having least total cost for this procedure}}$$

Because it is a typical archival practice to maintain systems files which contain a definition of all available variables, one testing condition applied equally to DATA-TEXT, SPSS and OSIRIS was that system files be used.

Figure 15.1 Descriptive Statistics Performance. (n = 1673, with 533 variables)
(^{O}Format statement size limit didn't permit running 400 variables)

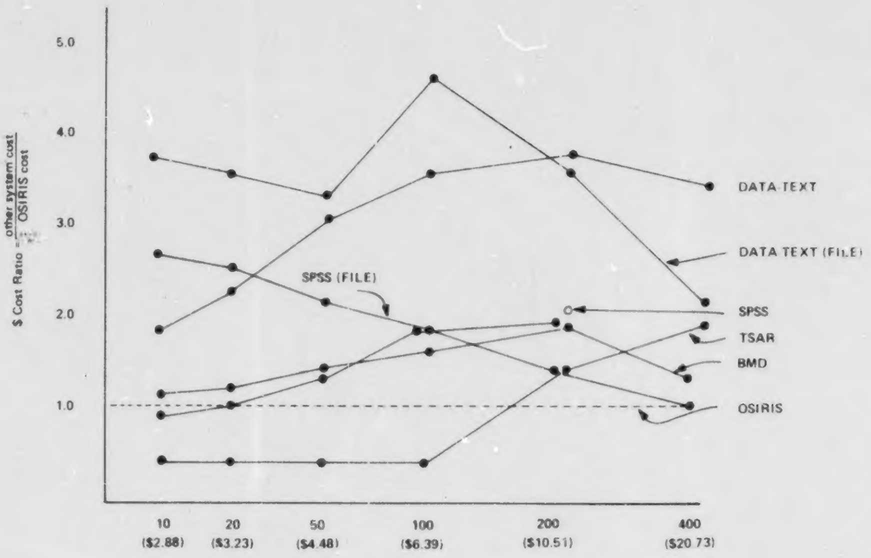


Figure 15.2 Bivariate Table Performance. (n = 1673, with 533 variables)
(°Lacks direct subsetting capability)
(*Tables based on a 25 percent subsetting of the original data)

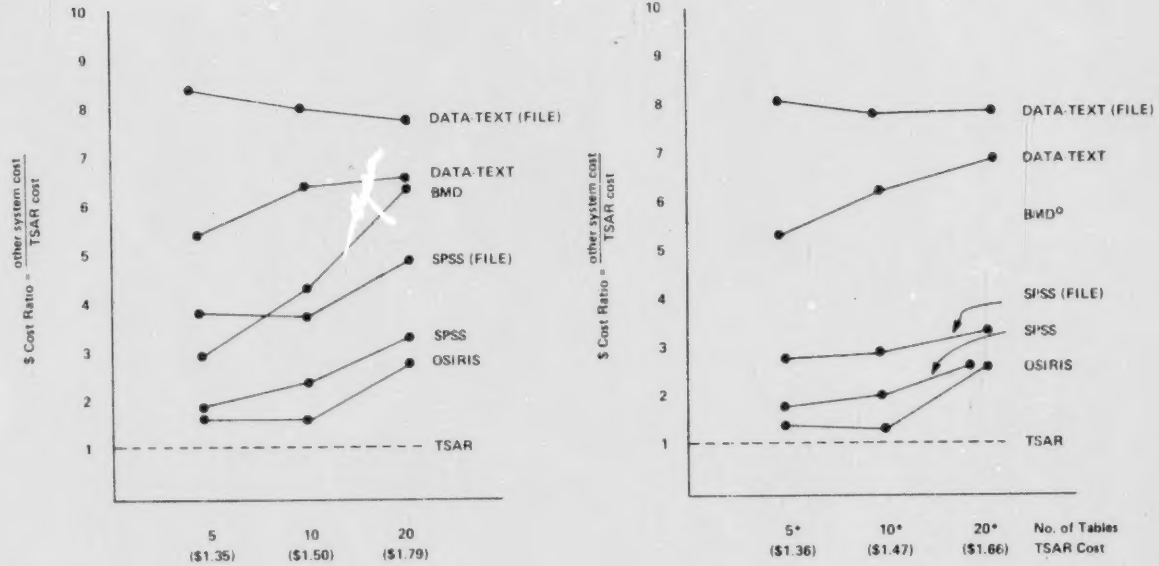


Figure 15.3 Pearson Correlation Performance. (n = 1673, with 533 variables) (°Insufficient capacity to run 30 x 30 matrix)

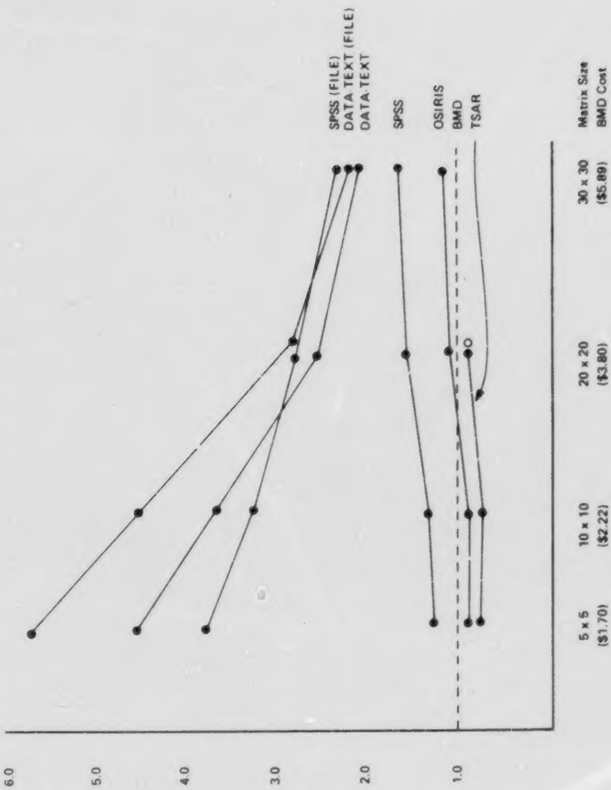


Figure 15.4 Stepwise Regression Performance. (n = 1673, with 533 variables)

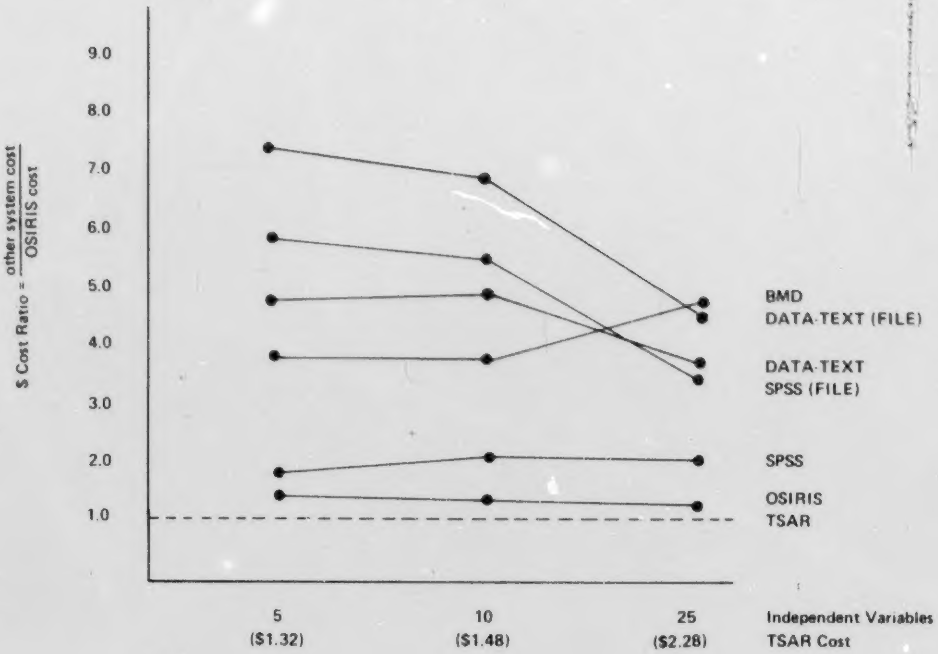


Figure 15.5 Factor Analysis Performance. (n = 1673, with 533 variables)
(^{o}Capacity Restriction)
(^{oo}No factor analysis program)

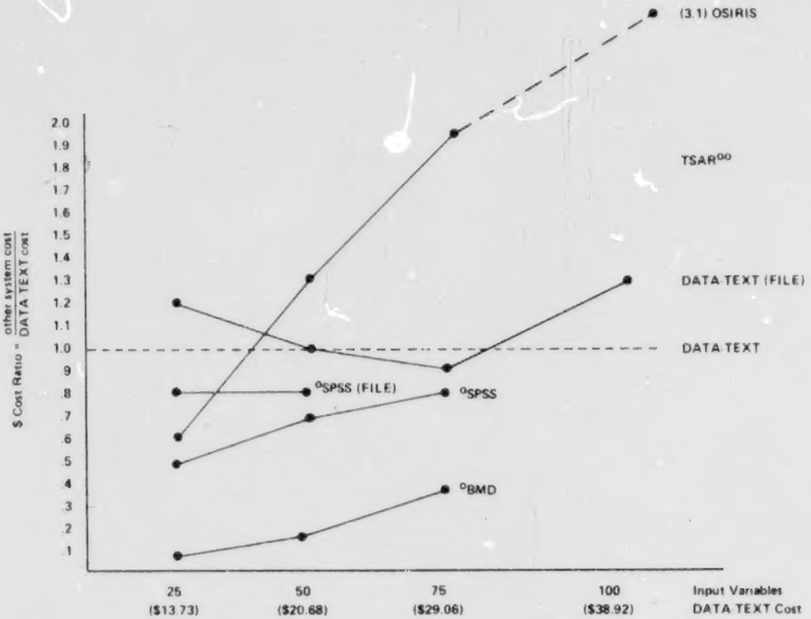
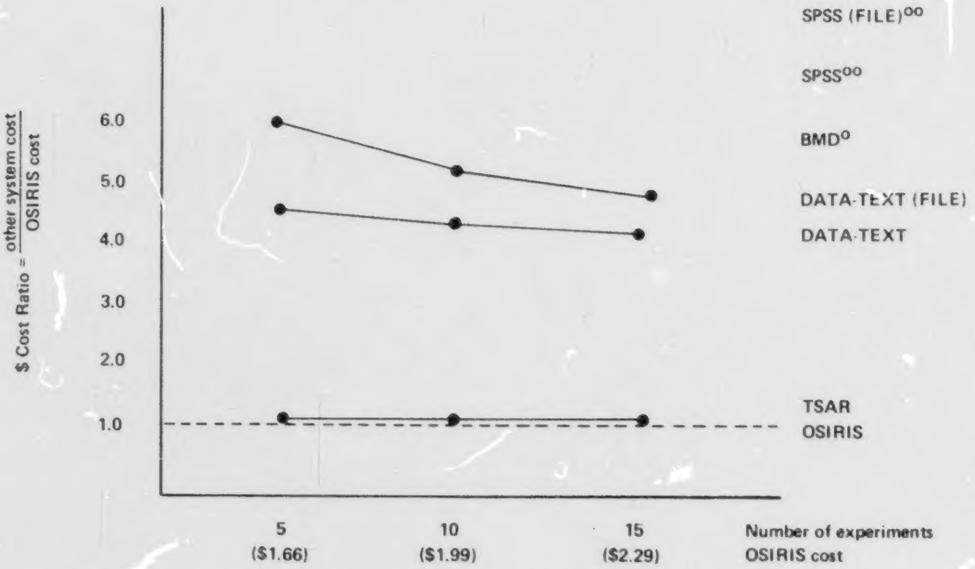


Figure 15.6 One-Way Analysis of Variance Performance. (n = 1673, with 533 variables)
(^{o}Input data format restriction)
(^{oo}No ANOVA program)



REFERENCES

1. This work was supported by the University of Connecticut Research Foundation, and carried out at the Social Science Data Center and the University Computer Center, University of Connecticut. These results are taken from a larger study: SLYSZ, WILLIAM D. "An Evaluation of Statistical Software In the Social Sciences," accepted for publication by the Communications of the ACM.

2. *BMD Biomedical Computer Programs,* W. J. Dixon, Ed., (University of California Press, Los Angeles) 1971.

3. Armor, David J.; and Arthur S. Couch. *The DATA-TEXT Primer,* (The Free Press, New York) in-press.

4. *OSIRIS II,* (Inter-University Consortium for Political Research, University of Michigan) 1972.

5. Nie, Norman, Dale H. Bent, and C. Hadlai Hull. *Statistical Package for the Social Sciences,* (McGraw-Hill, Inc., New York) 1970.

6. *TeleStorage and Retrieval System, Users Manual,* (Computation Center, Duke University) 1973.

Special thanks are extended to Everett Ladd, Jr., Director, Social Science Data Center, Bernard Lovell, Assistant Director, University Computer Center, University of Connecticut and to Eleanor Wilcox, Margaret Pyne and William Gammell, Social Science Data Center.

Software Compatibility in Social Science Computing

by Judith Rowe
Princeton University

Some years ago Princeton University held a conference on computing. One of the major concerns expressed by speakers at that meeting related to software compatibility. A FORTRAN II program written for an IBM 7094 would not run on a *UNIVAC* machine, etc. However, looking into the future one of the speakers saw the handwriting on the wall. He felt that these problems were, even at that time, capable of solution. The real concern he felt was data compatibility. Since the typical computer scientist knows little about data and cares less, the comment was lost in the relative merits of programming language, hardware and operating systems.

A year or two later in Boston at, what I believe was the first SIGSOC meeting, a speaker from MIT bemoaned the fact that our data sets couldn't talk to each other or more specifically files produced by one package were unausable by any other package. The problem still exists today but has finally been recognized. The earlier problem of software which runs on different computers has been solved most elegantly – by P-STAT. However, all of the major packages do now come in a variety of sizes and versions and can be run on a number of different computers. The elegance of P-STAT results from the fact that all of the versions are in effect the children of one set of parents and as yet none have left home. SPSS, DATA-TEXT and OSIRIS on the other hand have all sent their children to be reared in foster homes and although heredity remains constant environment tends to be an unknown quantity. The problem can be solved. One can produce a version of any of the social science packages which will run on any computer. The know-how is there and the principle of writing in higher level languages, even at the sacrifice of machine efficiency, has been accepted in most instances. It's merely a question of time and money. No one wants to live through another DATA-TEXT episode.

In the next circus ring, however, there has been a lot more footdragging and the reason is obvious. Why should the writer of package A spend his time making it easy for his users to read a data file produced by package B if the writers of package B won't do the same and in fact won't even reveal what their data files look like? The effort to convert a package to another machine produces an immediate payoff, especially for the producer who gets there first; efforts to implement a system of data file compatibility benefit no one, except the social science research community.

After the 7094 left Princeton (therefore post DATA-TEXT and pre-SPSS) many of the staff at the University thought P-STAT which was home-grown and potentially responsive to Princeton's needs could be modified somewhat and could be all things to all people. In theory, it is a good idea, but in practice at least at Princton it doesn't work. In addition to P-STAT, the University now offers users: 2 sizes of SPSS; 2 sizes of DATA-TEXT; an old and a new OSIRIS; and the BIOMEDS. Offering a diversity of packages produces more than a minimal strain on the clinicians and on the SSUS programmers. but if Princeton is to continue to be a user-oriented center there is no choice. Just as no one has come up with an ideal husband, no one has come up with an ideal package, and furthermore no one ever will. People choose packages with the same degree of emotion as they choose spouses and they stick to them with even greater tenacity, but perhaps for the same reasons.

If this seems unfair, let's try another analogy. There are people who walk, people who ride bicycles, drive cars, travel on buses, trains, ships, and airplanes and some people who do all of these things. To date no one has suggested that one mode of transportation will satisfy all needs. Some people like to travel slowly and cheaply and see everything along the way. Others care only for speed. Some people want to do their own driving while others want to leave the driving to Greyhound.

Because a similar diversity of tastes is evident among users of social science data, there is no reasonable hope that users can be limited to one package. Since users have to live in a multiple package world, it is imperative to increase the ease with which one can move data files from one package to another. If this is not done, every package will end up needlessly large, cumberson, and difficult to use in an effort to include even the most obscure of capabilities. Some steps have been taken in this direction. OSIRIS can now be read by SPSS and P-STAT.

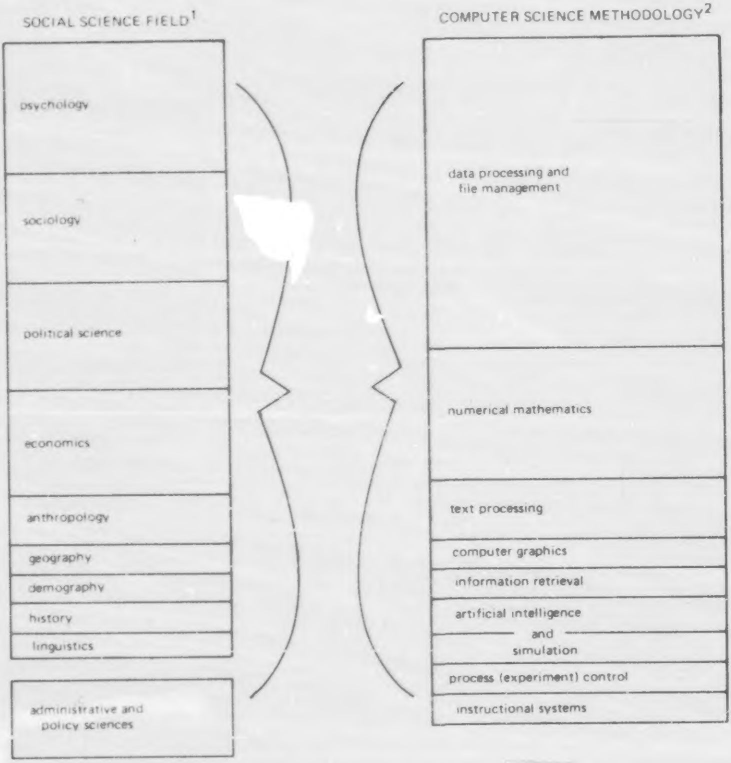
Reducing Incompatibilities in Social Science Software and Data: A Report on the SIGSOC Workshop

by Ronald E. Anderson
University of Minnesota

The development of computer software tailored to social science research is less than two decades old, yet in a very short time numerous programs, packages, and computer-based data sets have been developed. Since there has been no coordinated international or national effort to pool resources or avoid duplication of effort, most activity in the social science computing area tends to be highly localized and relatively difficult to export. There are exceptions but the field is best characterized by various types of incompatibilities. The dilemma occurs in part because of the diversity within social science research. Figure 15.7 which lists the social science fields and computing methodologies, reveals the variety of activities which force heterogeneity upon social science computing. Since developments in one area do not contribute to all other areas, incompatibilities occur. For example, content analysis does not help out the laboratory automation effort. However, since social science computing is still rather small, researchers can afford to accept some incompatibilities while jointly working for common causes.

Certain incompatibilities are detrimental. Cline[1] identified compatibility as one of the foremost problems facing social science computing. Noting that most computer programs and social data files cannot be exchanged without considerable effort among social scientists, he elaborated the problem as follows: "Today, the problem of exportability

Figure 15.7 Social Science Fields and Computer Science Methodologies


SOCIAL SCIENCE FIELD[1]
psychology
sociology
political science
economics
anthropology
geography
demography
history
linguistics
administrative and policy sciences
COMPUTER SCIENCE METHODOLOGY[2]
data processing and file management
numerical mathematics
text processing
computer graphics
information retrieval
artificial intelligence
and
simulation
process (experiment) control
instructional systems

Size of box indicates estimated amount of computing currently done in each field or the extent to which each computing methodology is used.

[1] List of fields is consistent with the definition of the social and behavioral sciences contained in the Report of the Special Commission on the Social Sciences of the National Science Board, *Knowledge Into Action: Improving the Nation's Use of the Social Sciences*. U. S. Government Printing Office, 1969.

[2] Categorization of computer science methodologies is taken from ACM Curriculum Committee on Computer Science, "Curriculum 68," *Communications of the ACM*, 11, 3 (March, 1968), p. 155.

has reached staggering proportions. Individuals find that many programs and data prepared for initial processing on one computer simply cannot be run at another computer center which uses exactly the same model and manufacturer. Local variations in peripheral devices and programming procedures make it extremely difficult and sometimes impossible to move from one computer to another.

Although one may place the blame upon the designers of the early computing systems for not having the foresight to think in terms of standardized data files and programming conventions, or one may cite the commercial manufacturers for not establishing industry-wide standards, the problem still remains; and social scientists cannot now realistically look to others for solutions. They must among themselves agree upon *common standards for program design and data file preparation* [emphasis, mine] which will minimize the problem of exportability. Granting agencies, both public and private, which support the development of social science computing can be of great assistance in insisting upon adherence to standards; but the original impetus for this movement must come from the social science community." In another paper Sadowsky[2], projecting into the future of social science computing, unintentionally supported Cline's position. Sadowsky identified "inadequate standards for data documentation and problems of data transfer" as among the most serious problems facing our rapidly expanding field.

These authors succintly express the consensus among social science computing specialists that serious effort should be directed toward reducing incompatibilities among computer programs and data sets. Cline goes on to argue that the impetus for this effort must initially come from the social science community itself. While few would disagree with this assessment, almost everyone has been waiting for someone else to work on the problem. Meanwhile many program and data sets have been generated with unnecessarily incompatible features. The SIGSOC workshop held June 1973 was the first public meeting designed to focus social scientists efforts toward resolving incompatibilities.

SIGSOC COMPATIBILITY WORKSHOP

Under the sponsorship of SIGSOC, a working session was planned and held during the National Computer Conference in New York City on June 6, 1973. Even though the meeting was not financially supported, over thirty persons attended. The attendance level and the intensity of participation clearly indicated that a timely, critical problem area had been identified. The format of the session was arranged to first characterize the existing situation within social science computing and data archives. The program is included in Appendix E. T. Dunnagan reported on CONDUIT recommendations dealing with technical transportation issues. The

CONDUIT project encountered surprising difficulty in simply moving a few data bases among five major academic computing institutions. Coover proposed a design solution for developing software for small computers, and made the effective argument that one loses the largest consumer sector if one design packages only for large computers. Klensin, never-the-less insisted that a "utility" approach was most desirable, offering the Consistent System of the Cambridge Project as an example. The Cambridge Project Consistent System is a large federation of packages that can interchange data with the aid of a data description scheme code. Buhler, who invented and cares for P-STAT, provided an excellent tutorial on the details of various FORTRAN implementation of P-STAT and how they are handled by a pre-processor. Steinbrenner reported on SPSS system progress and the various routines used to interchange data files with other major systems such as OSIRIS and DATATEXT. Marks reported on the very large data dissemination activity of ICPR at the University of Michigan. In particular he stressed the likelihood of finding environments with highly unique hardware and software. Shanks provided a useful history of the machine readable codebook issue, and characterized the codebook as the key to compatability of social data and program packages.

After the presentations, workshop participants were asked to complete a questionnaire on compatibility in social science computing. Responses to the questions were quickly tallied and reported to the group before the end of the meeting. Figure 15.8 contains both the questions and the frequency distributions of responses. Questions cover three compatibility areas: 1) packages; 2) data distribution, and 3) social organization. In response to questions one and seven which focused on package issues, an overwhelming majority agreed that a task force should be formed to construct guidelines for future data analysis program packages and modifications to existing ones. In addition, most agreed that those responsible for package distribution be encouraged to supply data conversion routines to handle files produced by other major packages. The consensus on data distribution center activities sought by questions eight and nine was slightly less extreme, but most agreed that a "definition" or guideline for data sets should be attempted and that is it probably feasible to do so. In addition, most favored the "national standard" rather than the alternative to "supply any of a variety of machine readable codebooks" as a requirement for all data distribution centers. With respect to the social organization of the social science computing and data dissemination sectors, respondents favored several organizational routes. The majority of responses favored formation of a consortium of social science data distribution centers, a national computing laboratory for the social sciences, and a clearinghouse for social science software. Later discussion revealed that many participants felt a clearinghouse like the one existing at

Figure 15.8 Tabulation of Answers to Compatibility Questions by 24 Attendees at SIGSOC Meeting, June 6, 1973

	Yes	No	?
1. Should a task force be formed to develop a set of guidelines for future social analysis packages and modifications to existing packages?	18	2	4
2. Do you favor the formation of a consortium of social science data distribution centers?	17	2	4
3. Do you favor the formation of a national computing laboratory for the social sciences?	13	4	7
4. Do you favor the establishment of a clearinghouse for social science software?	14	5	4
5. Do you think it would be possible to define a data set that could be utilized in nearly all social data analysis packages and that nearly any general purpose computer with a tape drive could produce?	14	5	4
6. Do you think such a data base definition should be attempted?	20	1	2
7. Do you favor the idea of encouraging all package developers/maintainers to supply routines which read or convert data files produced by other major packages?	18	3	2
8. Do you favor the idea of encouraging all data distribution centers to supply a machine readable codebook conforming to a national standard?	15	4	4
9. Do you favor the idea of encouraging all data distribution centers to supply any of a variety of machine readable codebooks conforming to any major package requirement?	6	11	6

the University of Wisconsin should be maintained.

Overall the results show that many knowledgeable persons, including persons with vested interests in packages and data archive centers, believe that incompatibility problems are sufficiently severe to warrant several different kinds of action. Furthermore, there seems to be a healthy optimism about the prospects for solving problems. Although there may have been a tendency for some to agree with almost anything proposed in the direction of compatibility, the pattern of responses indicates discriminating respondents who believe that large steps should be tried. The discussions that took place during the second portion of the working session were generally consistent with these questionnaire findings. It also became apparent from the discussion that most felt that the first order of activity should be meetings where attention is given to interpackage communication. With relatively little effort, techniques could be developed so that all the major packages can easily share data files.

PROPOSED WORKSHOPS

In order to take advantage of the momentum of attention to compatibility problems, it is imperative that a series of additional meetings be held and that these meetings be used to prepare a series of papers that are widely circulated. In addition a working structure of groups and committees should be set up to investigate issues on a long term basis. Workshops should be held in three areas:

- Data analysis software guidelines
- Codebook guidelines
- Data set transport guidelines

In addition, a wrapping-up workshop should be held that gives attention to organizational issues such as clearinghouses. The workshops should be open meetings, but formal presentations and papers would be solicited. Hopefully, those preparing acceptable papers would receive expenses and an honorarium. The productivity of the workshops should be combined into a book or extensive report and receive wide dissemination.

For the Data Analysis Software Guidelines Workshop the first order of activity would be a preliminary meeting of systems people representing each of the major packages. These persons would pool information and construct interface routines for accessing internal data fields prepared by and for other packages. A second preliminary meeting should take place to specify the needs, priorities, and possibilities for small installations, especially small liberal arts colleges. The workshop goal would not be to design an ideal language nor to bless any existing ones. Rather, it would be to identify preferred language features and programming techniques.

The Workshop Codebook Guidelines is probably best broken into at least two sequential meetings. The first meeting could attempt to bring together all necessary descriptive information to identify both the issues and the alternative actions. The second meeting would react to specific proposals.

A structure similar to that proposed for the codebook problem might be used for the workshop on Data Set Transport Guidelines. Attention should be given to issues such as documentation attributes, certification, coding systems, and exchange techniques.

THE NEXT STEP: ESTABLISHING A PERMANENT PROBLEM-SOLVING STRUCTURE

The SIGSOC session found that there is substantial pressure for guidelines or standards within the social science computing field. One of the key problems of standardization is timing; standards can easily appear too early or too late. Premature standardization may attack the wrong issues and freeze things that are not sufficiently stable. However, if standardization is delayed, too many forms appear in the marketplace which generates vested interests and unwillingness to accept the standards. Premature standards increase the risk of stifling new and creative ideas. Although it is impossible to eliminate risk, it is possible to minimize that risk by delaying standardization and by allowing some diversity to exist.

The best example we have in the United States of effective standardization is the ANSI (formerly ASA and USASI) organization, which is the authority for industrial standards in the United States. ANSI establishes working groups in particular areas; these working groups in turn can create technical subcommittees. The working groups investigate an area and decide whether or not to develop or modify a standard. If they draft a proposed standard, ANSI releases the proposed standard for a trial period after which acceptance or rejection is voted upon. Some aspects of the ANSI structure and experience from other standard organizations can be usefully borrowed in planning a strategy of attack upon incompatibilities in social data and program packages. A series of working groups can be established to investigate particular problem areas, and these working groups can report to a larger professional organization such as SIGSOC. Contacts and working arrangements can be made with other professional organizations through the parent group such as SIGSOC.

The proposed workshops are important because they would: 1) specify needs and possible solutions; 2) foster new informal contacts reducing unnecessary gaps among persons working on similar problems; 3) lay groundwork for an ongoing formal structure to deal with issues related to standardization; and most importantly, 4) actually develop new solutions, technical and otherwise, to exchange data. Incompatibility is a serious problem and there are people waiting to spend some time working on the problem.

REFERENCES

1. Cline, Hugh F. "Computer Instruction for Scholars in the Humanities," *Computers and the Humanities*, Vol. 3, No. i (Sept. 1968) p. 31-40.

2. Sadowsky, George "Future Developments in Social Science Computing," *Proceedings of the AFIPS Spring Joint Computer Conference* (AFIPS Press, Montvale, N.J.) 1972 Vol. 40 p. 875-83.

BIBLIOGRAPHY

Allerbeck, Klaus R. "Data Analysis Systems: A User's Point of View," *Social Science Information* Vol. 10, No. iii, (June 1971) p. 23-25.

"A report on Workshop on Computer Programming Systems for the Social Sciences," *Social Science Information* Vol. 10, No. iii, (June 1971) p. 39-46.

Alt, F. L., "The Standardization of Programming Languages", *Proceedings of the ACM 19th National Conference* (Association for Computing Machinery, New York) 1964 p. B.2-1-B.2-6.

Anderson, Ronald E., "A Bibliography of Social Science Computing," *Computing Reviews* (to appear).

Anderson, Ronald E. and Coover, Edwin R., "Wrapping up the Package: Critical Thoughts on Applications Software for Social Data Analysis," *Computers and the Humanities* Vol. 7, No. 2 (Nov. 1972), p. 81-92.

Armor, David J. "Developments in Data Analysis Systems for the Social Sciences," *Social Science Information* Vol. 3 (June 1970) p. 145-46.

Bisco, Ralph L. *Data Bases, Computers and the Social Sciences.* Information Science Series (Wiley, New York) 1970

Castellan, N. John Jr., "Laboratory Programming Languages and Standardization," *Behavior Research Methods and Instrumentation*, Vol. 5, No. 2 (1973), p. 249-52.

Cline, Hugh F. "Social Science Computing - 1967-1972," *Proceedings of the AFIPS Spring Joint Computer Conference* Vol. 40 (AFIPS Press, Montvale, N.J.) 1972 p. 875-83.

Goodstat, P. B., "Standards in Data Processing," *Data Processing Magazine*, Vol. 9, No. 3 (Mar. 1967) p. 22-25.

Meyers, Edmund. *Time-sharing Computation in the Social Sciences*, (Prentice-Hall, Englewood Cliffs, N.J.) 1973.

Simpson, Dan. "Future Trends in High-Level Languages," *Data Processing* (Mar.-Apr., 1973), p 89-91.

Steel, T. B. Jr., "Standards for Computers and Information Processing," *Advances in Computers*, F. L. Alt and M. Rubinoff, eds. (Academic Press, New York) 1967, Vol. 8 p 103-52.

PART V

COMPUTERS IN RESEARCH

Chapter 16

Computers in Research: An Overview

by Douwe B. Yntema
M.I.T. and Harvard University

The most striking impression one gets from these papers is that the use of computers in research has grown up.

Think what these sessions would have been like if this meeting had been held even a few short years ago. We would have spent most of our time talking enthusiastically about the potential of computers in scientific and scholarly research — talking about the good things computers were going to do for us someday. Somebody would have given a paper in which he said, in effect, "Look, it *is* possible to use a computer to control laboratory apparatus. I've shown it's feasible, and I almost have it working; and when I do, it will have a big effect on research in our laboratory." Some social scientist would have said, "It *is* possible to contrive a convenient, handy set of routines into which data from an attitude survey can be fed — routines that will permit an investigator to examine the data and understand their implications without having to write any programs himself. I have such a package almost working, and when I do it will have a big effect on social research." Even the number-crunchers, who were much farther advanced in their reliance on computers, were excited about bigger, cheaper machines that would let large numbers of researchers do computations that had been on the edge of practicality. We were going through an era of what I have been tempted to call gee-whiz salesmanship, a time when many people felt obliged to put large amounts of energy and enthusiasm into persuading each other, and the rest of the world, that computers could have an immense effect on scientific and scholarly research. It is rather startling to remember that all this was less than ten years ago.

Today the atmosphere is very different. The computer is evidently accepted on the university campus as a necessary tool of scientific and scholarly research, and the papers that were given at this meeting reflect that maturity. A healthy number of them were still concerned with new developments that are coming over the horizon, but many of them were concerned with the kinds of questions that maturity brings - questions about efficient arrangements for providing the computing services on which research has come to depend.

It was significant that even in sessions on the use of computers in research the authors of some of the papers used concepts from business and economics -- channels of distribution, brokerage, capital investment in databases -- and there was even discussion of the wisdom of buying from the growing information industry access to some of the information the academic world wants. From another point of view, there was a continuing concern about sharing resources in such a way that busy researchers in one place need not waste their time duplicating what has been done elsewhere.

It was also significant that in his keynote talk Ashenhurst had emphasized that the problem of giving researchers access to the computing services they need (in his case, researchers with minicomputers in their laboratories) is as much a problem of providing the proper people as it is a problem of providing the software, the communications interfaces, storage, and so on. This did indeed turn out to be a keynote: it was repeated again and again, sometimes with emphasis and sometimes only in passing, thoughout these sessions. Dedicated people (note the difference between a dedicated person and a dedicated machine) are a crucial component of arrangements for mature use of computers in research. The people, and the organizational structure that permits them to serve as a channel of communication, are as important as the hardware, the programs, and the databases.

Some areas of research are a long way ahead of others in this respect, and oddly enough, the list of participants in this meeting shows which areas have progressed farthest. There is, for example, a conspicuous scarcity of physicists whose research demands large amounts of raw computing power. A few years ago they would have been here; this fall they are not. Does this mean they are no longer using computers? Of course not. It means that on most campuses the computation center now includes people who have learned how to act as the agents for such users. Those people are here, and so the physicist does not feel constrained to come himself. On the other hand, the participants have included behavioral scientists, a physician, a surprising number of humanists, and a number of librarians. The lively discussions among them have been valuable in their own right, but from another point of view, the fact that those people have been eager to be here themselves implies that their disciplines could still use more help from kind of organizational structure and the dedicated people to which Ashehurst referred.

Asking what subjects did not get emphasized is almost as interesting as asking what subjects did. In the first place, it is remarkable that there were almost no discussion of plans to make use of networks of computers. The reason is fairly obvious. The whole network situation has been so fluid for the last six months that planning for the use of networks has to wait until things settle down.

Second, there was a curious lack of discussion about plans for regional or national computing centers specialized to fill the needs of particular disciplines. The development of such centers has already begun, and there will surely be more of them. Again, the uncertainty about networks has clearly been an obstacle to concrete plans about the way these centers will be used.

Third, it is surprising that so little was said about standardization. Standards for programs are almost a necessity if the programs are to be easy to transport, and standards for databases are crucial if the data are to be available for use with other programs. Given the continued interest in the sharing of resources, more talk about standardization might have been expected. I have mixed feelings about this subject. The Consistent System that the Cambridge Project is building is, from one point of view, an experiment in standardization. The practical difficulties are formidable, as those of us who have been involved in the project can testify. Perhaps postponing standardization until you are really driven to it, as we were, is actually a good idea.

But those are questions for the future. The main impression I got from these sessions is that we have come a long, long way in a few short years. The use of computers in academic research has grown up. We have taken off our short pants and come of age.

Chapter 17

Information Services for Research

New Channels of Distribution in the Information Industry

by William R. Nugent
Library of Congress

The growing maturity of the information industry is causing a shift from an initial preoccupation with technical processes to a concern with user needs, which is another name for marketing.

A trend to functional differentiation is apparent as organizations active in the information industry develop that are similar in function to the manufacturers, retailers, and brokers historically associated with tangible goods.

These new analog organizations in the information industry have differing functional roles, but most significantly they represent the formation of new distribution channels in the information marketplace that will have the greatest impact on the information industry in the next decade. In the 1930's[1], emphasis on distribution transformed the hard goods industries and created major benefits for the consuming public. It is highly likely that similar results will obtain in the new and post-industrial information business once more effective means of distribution are implemented.

While the non-profit sector has thus far played a major role in developing and operating research information systems, many of these organizations have shown a reluctance to develop marketing and distribution channels. It appears now that these organizations must begin to realize that even the non profit sector has a need, if not an obligation, to participate in marketing activities.[2] The alternative will be to become a user rather than a supplier of research information services. Even government information services that have normally been extremely conservative in marketing, are becoming more user and marketing oriented. A small but significant example is the recent decision of the National Technical Information Service, NTIS, to accept American Express charges.[3] The operative principle here is simply that of making the product or service easier to buy. This is one major function of improved distribution channels.

EDUCOM and the researchers, librarians, and educators that have participated in EDUCOM affairs in recent years have given much attention to the development of information networks. However, networking is merely a technical and organizational means to a marketing end. A network first and foremost is, or should be, a distribution channel in the marketing mechanism of the information business. Networks that have been such distribution channels have succeeded; those that have not emphasized marketing functions have remained underutilized technological curiosities.

More than a century ago, at the first faint dawn of the information age when telegraphy was just beginning, Henry David Thoreau observed: "We are in a great haste to build a magnetic telegraph between Maine and Texas, but who knows whether Maine will have anything to say to Texas?" While not normally noted as a business philosopher, Thoreau was correct in his assessment, that a need was not evident and a market had not been developed. Early telegraph networks were little-used transmission networks with little information of consequence flowing in them. Remote chess matches were frequent, as were test transmissions like "I'm OK – You're OK". It is interesting to note the similarity to the games and tests that constituted the early traffic in the ARPA network.

Once marketing was applied to these telegraph channels, they became distribution channels of high utility and high demand. Computer scientists can learn from telegraphy experience. One should attenuate his concern with networks qua networks, and concentrate rather on the need-filling functions of these networks. This is already happening in certain sectors of the information business, and it is probably accurate to say that the 70's will be to the new information economy as the 30's were to the older manufactured goods economy. Both periods could correctly be called "the age of distribution." The following examples of this new age describe the functional differentiation taking place in the information economy, and the emergence of manufacturers, retailers, and brokers.

INFORMATION MANUFACTURING

In the information industry one of the major manufacturing activities is the creation of machine-readable data bases. Other manufacturing concerns are multi-media publishing and SDI. Data base manufacturing is quite different from the manufacturing of tangible goods:

- The tools and techniques of production are, in general, not specific to a particular product.
- Costs are lumped in the first finished product off the manufacturing line, and the marginal costs of product replication are relatively negligible.
- Because the information product manufacturer can produce anything, with respect to content, he rarely has a proprietary claim on any one product, yet he can be competitive in highly diverse disciplines.
- Because his skills are process oriented rather than product oriented, the information manufacturer can be, and frequently is, an independent service firm with little or no ownership position in the source data.

Among the problems of information manufacturing are those of productivity, product life, and the establishment of corporate identity. Many aspects of data base manufacturing are labor-intensive and subject to wide ranges in worker productivity, somewhat like the programming field. The information speedup, particularly in research information, leads to a decreasing half-life of product utility, and requires faster amortization by users of product cost. The corporate identity problem arises because the information manufacturer works for information owners who, of course, place their labels on the manufactured products.

However, marketing solutions exist for at least the latter two problems. The reduced half-life problem requires greater emphasis on timely information products that communicate current news rather than providing access to archives. The corporate identity problem requires the creation of "house brands" of proprietary information products, using, for example, licensing arrangements with information owners permitting the creation of secondary and tertiary information by-products.

With a process orientation and consequent rapid flexibility, the information manufacturer has one major advantage over his counterpart in hard goods manufacture. When new fields of research are on the ascendancy, such as the fields of cellular biology and energy conversion appear to be today, it is easy for the experienced information manufacturer to use his tested and proven processes of data base manufacture to serve the new fields.

INFORMATION RETAILING

Retailing in the information business is like traditional retailing in that it provides time and place utility and information about the products and services of remote suppliers. The information retailer also provides the classic function of bulk-breaking: buying in quantity and selling in small lots. However, simple similarities end here. Retailing in the information business is a singular phenomenon that has no counterpart in the retailing of manufactured goods. Due its computer orientation, a distinguishing characteristic of infomation retailing is high fixed costs.

Retailers of manufactured goods can generally be characterized by financial measures of ratio analysis. Examining four ratios that concern sales, capital, and assets, one finds that the new "information stores" are a hybrid, in some new dimension, of jewelry stores and supermarkets. The continuing formal industry classification of the latter enterprises is "grocery stores," although its embodiments are, in vast majority, supermarkets. For the four ratios considered, jewelry stores, being inventory intensive, exhibit low ratios, whereas supermarkets, being asset intensive, exhibit high ratios. Information retailers show ratios representing apparently inconsistent extremes. Figure 17.1 exhibits the extreme ratio analysis ranges in the retail business of jewelry stores and supermarkets, and shows the directional trend of information retailers.

The information retailer, like the jewelry store, will tend to have a very low ratio of net sales to net working capital due to the high level of working capital necessary to support the computer facility and staff. In the case of net sales to inventory, one would expect the information retailer to resemble the supermarket with a very high ratio that reflects turnover and essentially zero inventory of products. The information retailer is more like the Tastee-Freeze or cotton-candy vendor, with a product made on demand with special machines mostly from air. For the ration of net sales to net worth, the information retailer is probably more like the jewelry store with very low ratios. This is expected because high working capital is required and capital is probably a major portion of net worth.

Finally, due to the expected use of leased assets, the information store exhibits a low ratio of fixed assets to tangible net worth, again like the jewelry store.

These comparisons suggest that the information retailer is handling a highly expensive product line, like the jewelry store, but hasn't yet learned how to distribute in quantity like the supermarket. Development of more effective distribution channels is the next major challenge to the information industry.

INFORMATION BROKERING

The information broker, like the retailer, provides time and place utility, but does not take title to the products or services sold. The most significant example of information brokering is that of the Northeast

Figure 17.1 Partial Ratio Analysis for Selected Retail Industries. Ratios are drawn from Dun's Review.[4] (1) Ratios for Jewelry Stores are based on 76 sample stores. (2) Ratios for Other Retailers are based on 2194 samples in 20 other retail categories. (3) Ratios for Grocery Stores are based on 134 samples.

Ratios	Information Stores	(1) Jewelry Stores	(2) Other Retailers	(3) Grocery Stores	Information Stores
Net Sales to Working Capital	probably lower	2.37	3.8 - 12.4	21.1	—
Net Sales to Inventory	—	2.7	3.8 - 10.9	16.1	probably higher
Net Sales to Net Worth	probably lower	2.02	2.7 - 8.8	10.3	—
Fixed Assets to Tangible Net Worth	probably lower	0.096	0.119 - 0.496	0.721	—

Academic Science Information Center (NASIC)[5], being developed by the New England Board of Higher Education with sponsorship from the National Science Foundation. Like brokers in other fields, the information broker is an aggregator of sources. One of his primary assets is timely knowledge of sources and prices. Unlike the broker in hard goods, 90 percent of whom work for sellers, the information broker works for buyers, and satisfies user demands rather than trying to sell the products of suppliers. Because the information broker is dealing in small lots rather than in wholesale lots, he must necessarily charge appreciably more than the broker who works for sellers.

Figure 17.2 lists typical, but by no means normative, commission structures of various brokerage operations. Prior to the emergence of the information broker the most recent addition to the group was the telecommunications broker who typically rents 240 khz supergroup channels and leases 4 khz voice channels to clients. Like the information retailer and the information broker, the telecommunications broker provides the services of bulk-breaking and short term user commitment. Because of the information broker's orientation to end-users and small lots, much higher commissions will be required for other brokers listed in Figure 17.2. Again one sees evidence that a new post-industrial organization cannot be measured by the parameters of earlier, and only superficially similar, enterprises.

The end-user orientation also means that the information broker is in several ways competing with the information retailer. One major and present problem in today's still-to-be-crystallized distribution channels of the information industry, is that of direct user access to all segments of the information industry. A user, presently, has the option of dealing with manufacturers, wholesalers, distributors, retailers, brokers, and other composite entities. Similarly, manufacturers can deal directly with retailers, wholesalers, and brokers. There is no clear distinction between retailing and wholesaling in the information business, since some organizations do both, and others even add a manufacturing component. This kind of chaos in the distribution of channels of manufactured tangible products, led to the trade wars and boycotts of the 1920's and 1930's, where combines of retailers and wholesalers often established boycotts of manufacturers who sold directly to retailers.

Arbitrage and futures brokering are two additional possibilities for the information broker. Arbitrage, in the money markets, consists of detecting differences in the prices of a given currency in different parts of the world and executing simultaneous buy and sell orders. In information arbitrage, users would place short term conditional buy orders for services at a specified maximum price and vendors would place short term conditional sell orders at a specified minimum price, according to the time of day, turnaround time required or offered, present computer capacity, and other

Figure 17.2 Typical Commission Structures of Brokers

Type of Broker	Typical Commission (percent of sales)
Stockbrokers (round lots)	1 – 3%
Hard Goods Brokers	2 – 3%
Other Investment Brokers	up to 5%
Real Estate Brokers	6%
Telecommunications Brokers	8%
Financial 'Finders' – Merger and Acquisitions Brokers	up to 10%
Information Brokers	substantially higher

factors. The broker would match the buy and sell offers as favorable combinations were found. In futures brokering the information broker would take long and short positions on future vendor resources as well as trading for customer accounts and for his own account. Since time, including computer time, is a highly perishable commodity, the information futures market would be highly volatile. While these speculations may appear remote, the information broker will, by one mechanism or another, perform the highly valuable function of finding a marketplace for suppliers' excess (and low marginal cost) capacity and providing best buy services for users.

FACTORS OF SUCCESS

This brief review of the new channels of distribution in the research information industry leads to several broad conclusions concerning the factors of success in the information industry.

First, research information services should seek to develop unique and distinctive services rather than competing one-for-one with other research information services. Since much source research information is copyrighted, there is a low cross-elasticity of demand among research information service organizations. The demand situation is similar to that of the general computer service industry.[6]

Second, like other operations with high fixed cost, research information retailers should capitalize on low marginal costs and seek to operate at near-capacity via improved distribution and marketing. A brokerage operation can be of particular benefit by aiding the retailer with distribution and marketing.

Third, information networks should be viewed as a means to achieve improvements in distribution through technological enticements.

Finally, in the post-industrial distribution age of the information business, one may expect that the manufacturers, retailers, and brokers will lose an initial similarity to the industrial age archetypes used for categorization purposes and begin to adopt newer market roles that will optimize their particular potentials in electronic distribution. However, as in the industrial age, distribution channels will both characterize and transform this new business.

REFERENCES

1. Joseph C. Palamountain, Jr., *The Politics of Distribution*, Greenwood Press Publishers, New York, 1968.

2. Benson P. Shapiro, "Marketing for Non-profit Organizations," *Harvard Business Review*, Sept-Oct 1973, Vol. 51, No. 5, p. 123-132.

3. "NTIS Honors American Express," *Electronics News*, Aug. 27, 1973, Vol. 18, No. 938, p. 21.

4. "The Ratios of Retailing," *Dun's Review*, Sept. 1973, Vol. 102, No. 3, p. 103-105.

5. David M. Wax and R. D. Morrison, Jr., "NASIC – A Regional Experiment in the Brokerage of Information Services," *Facts and Futures* (EDUCOM, Princeton, N.J.) 1974.

6. Richard S. Bower, "Market Changes in the Computer Service Industry," *The Bell Journal of Economics and Management Science*, Vol. 4, No. 2, 1973, p. 539-590.

NASIC:
A Regional Experiment in the Brokerage of Information Services

by David M. Wax and
R. D. Morrison, Jr.
New England Board of Higher Education

The Northeast Academic Science Information Center (NASIC), a program of the New England Board of Higher Education, represents an attempt to apply on a regional basis the concept of brokerage or wholesaling to the provision of computer-based information services. Through intense promotion and active marketing of services and through training of Information Service Librarians on the campuses of the major academic institutions, NASIC aims to create, tap and serve the substantial market of researchers in the Northeast who are in need of continuous and comprehensive information. Through aggregation of user demand and negotiation of bulk purchase contracts with multiple suppliers, NASIC intends to serve the research community of the Northeast at a reasonable cost while assuring its own continued viability as a self-supporting organization. Finally, through utilization of available computing capacity in the Northeast and an existing communications network in the region, NASIC ultimately aims to become an efficient, self-supporting supplier of low-cost services particularly appropriate to the needs of the region's research community. These goals and the efficiencies and economies that they imply are the basis for the decision by the Office of Science Information Service of the National Science Foundation to underwrite the development of NASIC, and for the support and enthusiasm for this cooperative effort by many of the principal universities of the Northeast.

FUNCTIONS OF NASIC

In more specific terms, NASIC will direct its efforts toward the fulfillment of eight basic functions:

- Market survey and analysis
- Promotion and marketing of computer-based information services
- Training of library personnel
- Aggregation of demand from multiple users
- Negotiation of contracts with suppliers at discount rates
- Development of document delivery capability
- Supplying information services directly
- Development of an efficient delivery system based on existing resources within the region

Each of these functions is directly tied to the viability of NASIC as an organization and the brokerage concept as an effective means of providing information services.

The activity involving the determination and evaluation of the market for computer-based information services underscores the basic principle underlying the NASIC approach to the provision of information services. NASIC is devoting a significant amount of its efforts and resources to an assessment of the specific information needs of the research community of the Northeast. To determine the nature and extent of the market, NASIC will utilize many of the techniques common to product introduction in the industrial and commercial environment, but definitely not typical of the academic environment. This analysis will enable NASIC to ascertain demand for particular information services and will allow the setting of prices necessary to assure NASIC's self-support.

In addition to the determination of the existing and potential market for computer-based information services, NASIC will attempt to create new users through an active campaign of advertising and promotion. Various methods and media including campus newspapers, bulletin boards and department meetings, will be utilized and evaluated for broadcast effectiveness and cost effectiveness. Recognizing that word-of-mouth is probably the least expensive and most effective means of publicizing the utility of information services, NASIC will concentrate on developing a cadre of satisfied users on every campus.

An important aspect of the NASIC concept is that computer-based information services constitute a logical extension of the information services and products traditionally provided by the college and university library. Within the library of each major research institution in the region, one or more staff members will be trained to function as Information Service Librarians (ISL's). With the understanding of the principles underlying the search techniques for data bases, with awareness of the content of the broad range of available data bases, and with the skills to develop and code effective search profiles, the ISL's will function as the interface between the supplier and users of information services so that the user can obtain fullest utility from existing information resources.

It will be the function of the ISL first to assist the researcher in the definition of his specific information requirement. Second, the ISL will have the responsibility to select the most appropriate source for the information required by the user. This decision involves not only the determination of appropriate data bases, but also an awareness of the search algorithms employed at the various information centers processing that data base to guarantee the most effective fulfillment of the particular requirements of the individual researcher. Third, the ISL will undertake the development of a search profile that will yield the greatest recall of citations and abstracts needed by the user while minimizing irrelevant and unnecessary references. Fourth, the ISL will be involved in the reviews of search output for relevance and completeness judgements. This review will also serve as the basis for determination of the desirability for the modification of the search profile in the case of continuing current awareness services. Finally, the ISL or a colleague in the library staff will assist in arrangements for document delivery related to the search output.

While the traditional role of the library is thus not modified by the introduction of computer-based information services, (a computer printed list of bibliographic citations and abstracts is not significantly different from a typed list) the manner and means by which services are provided by the library to the user will be changed. For this reason, an important component of the NASIC program will be the orientation of the academic library community to the utility of computer-based information services and the training of library personnel to enable them to acquire the knowledge and skills necessary to function as effective Information Service Librarians.

Another important aspect of NASIC's brokerage role will be the aggregation of demand from a multiplicity of users dispersed over a large geographic region. The economies of scale in the provision of such services are substantial, and only through aggregation of demand can the real costs of these services be kept at reasonable and marketable rates. The true costs of providing these services on a single campus, particularly in small or medium-sized universities, are very high and, in most cases prohibitive. Through regional aggregation of demand, the cost components including hardware, data bases, technical staff input, and administrative and management resources can be prorated over a large number of searchers, with the add on to any individual search kept relatively low.

Further economies of scale and competition can be generated through the negotiation of contracts with a range of suppliers at discount rates.

Because of the existence of competitive suppliers for many of the relevant services and because most of these suppliers have capacity to provide significantly greater quantities of services at only marginal increases in real costs, the discounts resulting from aggregation of demand should in many cases be sufficient to cover the entirety of the NASIC overhead costs. Finally, NASIC's activities as a broker and aggregator of demand will enable NASIC to provide the research community of the Northeast with a much greater variety of services than would otherwise be available.

The basic assumption underlying this brokerage experiment is that it is inconceivable that a large number of universities in the Northeast would be willing to underwrite on their own the provision of a broad range of computer-based information services. And if one or more institutions did make such marginal services available, it would be impossible for them to provide the great variety that NASIC as a broker and aggregator of demand could offer to all of the researchers throughout the entire region.

The provision of bibliographic and abstract services at a reasonable cost, while in itself a worthy goal, does not constitute the sole objective of NASIC. The user often requires services beyond a listing of citations and a collection of abstracts; in many cases, document delivery is an essential component of total service to the user. Through its close ties to the academic libraries of the region, NASIC will work toward the development of systems for delivery of the necessary hard copy to the user, whether it take the form of reprints, photocopies, microfiche or monographs. NASIC's link to the computer-based library support system of the New England Library Information Network (NELINET) and the ultimate utilization of that system for serials control, will play a significant role in the development of a viable hard copy delivery capability.

NASIC also recognizes that there is a broad range of potential information services that are not now available anywhere, and it is likely that NASIC will become a direct supplier of such services should a market for them exist. Activities in this area would include the development by the New England Board of Higher Education of data bases that would have national or particular regional utility. Such efforts might also include the initiation of a newsletter or adoption of other means of communication to keep the research community of the Northeast current on new developments related to information in their fields of interest.

Finally, NASIC sees as one of its primary functions the utilization of existing resources within the region as a basis for efficient delivery of information services. Reference has already been made to potential use of the NELINET system to assist in an economical document delivery capability. It is also planned that use will be made of the existing general computation computer network of the New England Regional Computing Program (NERComP) as a communications network for at least part of the delivery process for computer-based information services. Should the utilization of the NERComP network prove effective, it is possible that the network might efficiently expand to include the entire Northeast region. Furthermore, if demand for information from one or more data bases warrants it, NASIC could begin to spin tapes on its own, utilizing existing hardware and technical resources within the region to do the processing and the NERComP network as a delivery system. Finally, as NASIC moves into an operational mode, it is likely that we will discover or modify other existing resources to add to the efficiency of the delivery system for computer-based information services.

NASIC: THE USER'S PERSPECTIVE

The primary goal of NASIC is to provide valuable information services to the bench level researcher in the Northeast. From the point of view of this user, NASIC must provide fast and economical access to a body of information that is of use to him. Thus, to meet its objectives NASIC must be able to provide information from a broad range of data bases, must be able to deliver information products that are timely, complete and relevant, and must be able to do this at a price that will make the purchase of this service desirable.

It is the operating plan of NASIC that the user perceive these computer-based information services as the logical extension of the information services presently provided by the university library. When he is in need of information, the user will arrange to discuss his particular requirements with the Information Services Librarian (ISL) resident at his own institution. Should the specific subject area be outside the range of competence of the local ISL(s) or, should the user be located at an institution too small to employ a trained ISL, the user can be directed to an information specialist at the NASIC offices or to an ISL at another university in the region who happens to have special competence and knowledge in the specific discipline in question.

The ISL, after discussing the research problem with the user, will utilize the knowledge he or she has acquired about the various data bases and search techniques available to determine the most appropriate means of providing the necessary services. A search profile will be developed and coded by the ISL and the search ordered under the auspices of NASIC. The mode (on line or off line) to be used in developing the profile and ordering the search will depend on a number of factors including the availability of processing alternatives, the time requirements of the user, and the price the user is willing to pay. Within a period of time ranging from a few minutes to a week or more, the output will arrive at the desk of the ISL, who will review it with the user for relevance and completeness. At the same time, the user will be informed as to the options for obtaining hard copy of those articles that are of particular value.

In most cases the user will pay for these services by charging them to his personal or grant account at the university. In some instances, departments might allocate funds for acquiring these services in their budgets, as support to the research of department members. In other situations, the college or university might choose to make these services available to all faculty members and students either free or on a token charge basis, with the service subsidy coming out of general library or academic budgets. But in all cases the user will appreciate that he is receiving a valuable service which has a real cost attached to it. And, the user will also be aware of the fact that the availability of this kind of information service was the result of a cooperative effort of his own institution and similar institutions operating under the auspices and guidance of NASIC.

THE BROKERAGE CONCEPT

NASIC reflects the intent of the New England Board of Higher Education, with the encouragement and financial assistance of the National Science Foundation, to provide a large research community with economical and effective access to a broad range of information services. By utilizing existing human and machine resources, both within and outside the Northeast region, NASIC hopes to demonstrate that these services can be made available without paying for the redundance of reinvention. While implementation of an effective brokerage operation is essential to the long range success of this experiment, of far greater importance is the demonstrated willingness of the region's institutions of higher education to recognize the need for cooperation for mutual benefit and to work diligently to achieve that cooperation.

With a valuable product, effective management and continuing inter-institutional cooperation, NASIC will be able to demonstrate that the brokerage concept can be applied to the provision of information services.

The Ohio State University
Mechanized Information Center

by Hugh C. Atkinson
The Ohio State University

The Mechanized Information Center, funded by the Office of Science Information Service of the National Science Foundation, provides current awareness and retrospective search of bibliographically oriented information. It is now preparing 12,000 bibliographies for 4,000 persons monthly using data bases which include:

- Pandex
- ERIC including Research in Educational and the Current Index to Journals in Education
- Psychological Abstracts (retrospective search only)
- Chemical Titles (current awareness only)
- Bibliography of Agriculture (current awareness only)
- Institute for Scientific Information: Source Data and Social Sciences data
- NTIS, formerly the U.S. Government Reserach and Development Reports
- Current Index to Conference Papers
- MARC tapes

To make the operation financially feasible, all retrospective searches are run on a particular day from data which is sequentially stored approximating a six-month period per single disk pack. Permanent personnel for the Mechanized Information Center are paid from the OSU Libraries budget, and the NSF funds cover the developmental costs. When NSF funding ends in February 1975, the Center will continue as a normal reference service. The annual cost to MIC to run an SDI service of 4,200 current profiles plus 6,000 retrospective searches is approximately $120,000 per year. The MIC file holds approximately 3,000,000 entries, and computer time is billed at standard rates. The cost for operational personnel that probably will remain after the NSF funding drops off will be about $65,000 per year, which is equivalent to about five reference librarians. Cost for the developmental staff has run about $100,000 per year for four years.

KEY FACTORS FOR SUCCESS

On a campus like OSU where our catalog access library circulation system contains 60,000 valid names of borrowers, the geographic location where one gets data is very important. Centralized libraries, whether in cities or on campus, break down. There is a real urge for departmental libraries. One rule of thumb is that a serious member of the faculty will be willing to walk as far for information as he does from his parking place to his office. A student walks as far as from his class to the student union. Because OSU has twenty-two libraries on the campus, the Mechanized Information Center's services will not be centralized when the developmental phase is over, but will be available at the reference desk of each library. In addition to locus, the next most important factor for successful information service is hours of service. Since information needs exist in about the same time frame as other human activity, most patrons gather information from 7:00 a.m. to midnight. The MIC will provide its profiling services to all the reference desks and through any other agency during these hours. Another key to success is the development of simple information profiling techniques which will allow persons at remote locations to use the system with a minimum of tutorial interaction. The MIC has opted for less precision and more coverage in designing its system.

ON-LINE CIRCULATION SYSTEM

The OSU Libraries has experimented for the last two years with an on-line catalog access circulation system that has 1,000,000 author and title short entries for the books held by the University Libraries. In response to a telephone inquiry, the OSU Libraries provides information on whether a particular book is in or out, and can charge out the book and mail it to the requestor if desired. All of the Libraries' holdings, comprising 2,700,000 volumes, are encompassed by that file. The Libraries will close its card catalog on July 4, 1976, as its part in the National Bicentennial Celebration.

Because The Ohio College Library Center has been requiring all of the current cataloging for the past two years to be in full MARC format, with the exception of non-Western alphabets, theses, and some other peculiar forms, the OSU Libraries may have full data for most items acquired after 1970 when the card catalog is closed. Since speed is so important, the Libraries may use indexes rather than the traditional library combination of the item description and the index itself. In machine systems, one can easily separate the indexing, the location, and the description.

SUMMARY

The Mechanized Information Center will use its experience to provide a blueprint and the necessary research to make the closing of the card catalog a reality. The primary need is to provide subject access to the OSU collections.

The overall objective of all these systems is to make the OSU collections as accessible as possible to all patrons.

Data Bases and Libraries

by G. A. Harrer
University of Florida, Gainesville

This paper outlines the philosophy and describes activities at the University of Florida regarding the role of data bases in the library. Some major problems are also noted. At the University of Florida libraries, far from being simply large warehouses storing the printed word, are viewed as vital units whose responsibility to the parent organization resembles that of the memory portion of the brain to the human body. Libraries should collect, organize, store, and when necessary, retrieve any information of value to the larger organization, be they universities or cities or whatever.

Data bases are the very life blood of libraries. In the earliest forms, data bases were recorded by means of wedge shaped impressions in soft clay or scratches on palm fronds. Through the years, major advances have been made in the symbols used to record knowledge and the media on which these symbols were preserved. However, because of the early development of the codex form of the book, libraries have been thought of as large book warehouses as the name "library" indicates. It is true that the major medium used to record man's knowledge is the printed page, and the book remains the most efficient method for storing information. It is also fair to say that librarians have had to deal with the intrusion of machinery between man and the printed page, if only in the form of a magnifying glass for many years. Storage of the printed page in the form of micro images has been commonplace and libraries are certainly faced with this imposition. Why is the use of microfilm or other machine felt by many to be an imposition? Perhaps, because the machine renders the collection unbrowseable. It is almost impossible to leaf through a microfilm collection, and a deck of punched cards carries the impediment to the extreme. Both require the use of an intermediate machine, and some indexing assistance.

The difficulty of information retrieval has been exacerbated by the size of libraries today. It's not possible to go to the stacks and browse through the shelves of any but the tiniest speciality and efficiently gain information. For minor specialities a good deal of the information on the data bases may well be organized in a remote part of the library. Recognizing the problem, libraries have spent a great deal of money to create indexes to the data base, card catalogs, but the card catalog and available bibliographies have become so large and complicated that the libraries have had to increase staff to assist the lay patron in locating required information. Thus the reference department has become the expert transfer medium in the interface between the seeker and the information. As libraries have progressed from the magnifying glass to the computer, the complexity of this interface has increased immensely and has, in some cases, been allowed to wag the dog. It is the responsibility of the library to connect the user with the information whether the information is stored on a strip of microfilm, a phono disc, a punched card, or a large reel of magnetic tape. It is the librarian's responsibility first to clarify the question, and then to know how to get at the required information.

The computer is not an end in itself. Just as librarians were long ago accused of fostering libraries for the sake of libraries and catalogers were alleged to love catalogs for their beautiful detail, today computer specialists are sometimes thought to glory in the precise complexity of the computer without sufficiently considering their utility in the service of the public. This we must carefully guard against.

The University of Florida Library was one of the first libraries to run experimental searches of the earliest MARC tapes from the Library of Congress. Later having obtained from the National Agricultural Library the first reel of magnetic tape which contained the agriculture bibliography, the library further experimented with information retrieval for scholars in the agriculture experiment stations. In 1968, the library obtained a budget line item for a systems analyst/computer specialist, and in 1969 brought to the University Mr. Robin Fearn in the position of Assistant Director for Systems.

ICCC AT FLORIDA

In 1970, the University of Florida agreed with the nine libraries of the State University system to jointly purchase a membership in the DUALabs

"Start" program in order to have available for the University System 1970 Census data in useable form. When the University of Florida formally became a member of the Start program, it agreed to represent the Florida State University system at training sessions, and to hold later its own training sessions in Florida, alerting people to the availability of data. The library further agreed to obtain any data required by the State University System Libraries from DUALabs. By early 1971, it became obvious that the demand for census data was not limited to the University faculties but extended throughout the state of Florida for planning groups, businesses, and state Government. Requests for census data became so numerous that the library was forced to take some formal steps. A program of Information for Campus, Community, and Commerce was established as an integral part of the University Library with office and operational space within the main library complex. The Library Systems and Data Processing Group provides design procedures, coordination, and computer program support for ICCC computer operation. Although the library is equipped with some data processing equipment, ICCC uses the campus IBM 370/65 computer for most processing, and has storage vaults now containing well over a thousand reels of computer processable information files. The objective of the ICCC was to expand and improve informational products and services of the University libraries to enhance the instructional research and service missions of the University of Florida and the State University System of Florida as a whole. Specific activities in which this group has been involved to date include: consulting and advising other university departments and organizations on problems of information storage and retrieval; investigating the nature and structure of a regional planning operation thesaurus; developing census and other social and economic data service to optimize the utilization of both computers and printed data; developing computerized bibliographic current awareness and retrospective search services; developing a State Union List of Serials with monthly updating capacity; developing a book catalog for the State University Extension Division Library; and developing cooperative programs to exchange software, data bases, and other information products and services. ICCC personnel are based primarily in the library. Three members of the library reference staff who were specifically trained in the contents and organization of the data files, have obtained training in the basic programming techniques and spend almost full time helping the public access ICCC files. Also in the library two other staff members who are full fledged computer programmers work full time on the ICCC System. Mr. Fearn the Assistant Director, supervises the total operation and acts as advisor when necessary. In the computing center, several personnel are assigned primarily to library needs but are not on our budget.

Within the ICCC, the Census Access Program was established as a major section to deal with the census. Because many requests for census data came from industry outside of Florida the Census Access Program obtained full files of all census data for the total United States. However, availability of full information only increased the demands. The Program has served hundreds of clients from almost every conceivable organization and has run workshops both in Gainesville and elsewhere in Florida to inform the public about the organization and contents of census data. ICCC continues to process this data as a major service at the library for what appears to be an ever increasing demand. Using National Agricultural Library "Cain" tapes, ICCC has made available to the agricultural community of Florida a vast amount of bibliographic information, and the libraries are currently providing SDI service for over 200 paying subscribers. The libraries are also prepared to do current awareness searches and retrospective searches of Mark II tapes, although at present there is a very small clientele for this service. By agreement with the University of Georgia computing center, the University of Florida libraries attempt to avoid any major duplication of data bases and process requests for Georgia which require the use of the University of Florida data bases.

PROBLEMS

A primary problem in data handling in the library is cost. Since libraries traditionally have not charged the local community for information, basic costs have simply been borne by the university. University subsidy to the library pays for large collections of data, and for staff to organize them, put them on the shelves, file them away, and provide reference service. To date University of Florida libraries have purchased data bases with library funds for book acquisition. Book funds are for the acquisition of information for the library's data base, and this is certainly information for the library's data base. However, the problem of paying for computer time remains. On many occasions, faculty pay for the computer time and the libraries provide all other services for information retrieval without charge. Because the Computing Center at the University of Florida allocates time to departments, and the departments allocate time to faculty, it is possible for the faculty members to get time on the computer and absorb the library's cost for computing service. Although this procedure is aggravating the auditors in the administration building because there is no way of showing the true cost for the type of service the library is giving, the libraries through ICCC have been doing it any way. If searches are processed for the state government, faculty with grants, or non-university clients, a minor service charge is made for the use of the data base. If services are provided for commercial firms who are interested in some of this data, the libraries charge whatever the traffic will bear, and have been able to recoop some expenditures. Any profit is used to subsidize experimental searches of data or to improve the data base. The problem of fairly allocating the cost remains. How does one charge? Who does one charge?

A second concern is access to the computer. There are times when library programs simply had to be delayed because of other demands on the computer. Some formal relationship will have to be established to give the library a priority with the computing center or other alternative.

Chapter 18

Resource Sharing for Libraries

Major Trends in Library Computerization

by Richard De Gennaro
University of Pennsylvania

Activities and trends in library automation can be grouped into three major categories. The first and most important one is the area of *cooperative computer-based networks or consortia.* The Ohio College Library Center (OCLC) which is by far the largest and most significant of these systems, provides as a main activity a very effective on-line shared cataloging system. OCLC has 50 active members in its parent Ohio network and several satellite networks in various stages of joining or participating including NELINET, FAUL, PRLC, UCL-PALINET, the Federal Libraries group and SOLINET. Several other cooperative projects also deserve some notice. The SPIRES-BALLOTS project at Stanford University which was started several years ago, has received over one million dollars in development funds from the U. S. Office of Education and the Council on Library Resources. BALLOTS is a bibliographical system for a single library in an on-line mode. The system in its present stage is proving to be rather expensive for Stanford to operate on its own, and efforts are being made to expand it to include five other libraries from the San Francisco Bay Area in a cooperative arrangement. While BALLOTS is a successful working system, its national significance has been limited because it appears to be locked into Stanford's unique computer environment and cannot easily be transferred or replicated elsewhere in its present form. However, Stanford is planning to make the necessary modifications to the system in the next year or two to overcome this limitation and to make it possible for BALLOTS to be brought up on standard IBM 370 operating system software. This would increase the transferrability of BALLOTS and probably also reduce its present high operating costs.

The Department of Higher Education of the State of New Jersey contracted with IBM to develop and implement CAPTAIN, a central bibliographic processing system based at Rutgers University, designed to serve the needs of all the units of the State University and the state colleges. The system has both on-line and batch processing features and is apparently designed to make it possible or necessary for much of the processing of materials to be centered at Rutgers. Because CAPTAIN is in the early implementation stages, it is too soon to assess its success and long-range significance.

IBM has another major library system called ELMS which was developed at the IBM Los Gatos facility between 1965 and 1972. ELMS was to be a comprehensive package system capable of doing most of a library's technical operations. Although it was implemented at the Los Gatos IBM library, the company has not tried to market it to individual libraries. Instead, IBM renamed the system "Library 370" and is trying to interest a group of libraries in taking it over along with the quite substantial expense that will be required to complete the development and make it work. The Illinois Educational Consortium for Computer Services has apparently agreed to be the first taker for "Library 370", but several others are needed before work on the system can begin. One can question the wisdom of continuing to put substantial resources into completing a system for the long future which was based on concepts of library automation and system design which may by now be somewhat out of date.

Other cooperative systems are in varying stages of development and implementation in Oklahoma, Maryland, Toronto, and elsewhere. These few examples show that the automation of the acquisitions and cataloging functions in libraries through cooperative systems is a rapidly growing trend. It is clearly easier, cheaper, and less risky for libraries to automate or get the products of automation by joining networks than by attempting to develop stand-alone systems. Much of the glory has gone out of pioneering computer systems, and it is now quite acceptable even for a major library to have no in-house automation program and staff.

A second major category and trend in library computerization can be characterized as *vendor-supplied systems and services*. This approach to automation has much in common with the cooperative approach. While it is not yet as widespread and significant as the cooperative approach, it is also growing rapidly and could in the long run surpass it in importance, particularly for the mass of ordinary libraries. If it is easier, cheaper, and less risky for a library to join a cooperative than to develop individualized systems, it may be even more advantageous for a library to become a customer of a vendor and buy or lease the automated systems or products that it needs. This approach has not been important in the past because few vendors had any tried and true systems or products to offer. However, a number of vendors have begun to offer systems and products to libraries recently.

Book jobbers such as Richard Abel Co., Baker & Taylor, and Bro-Dart use computer systems for internal operations and are, therefore, in a favorable position to offer computer-produced cards, lists, and other computer-based services. Eventually, the customer library may be able to tie directly into the jobber's data base and system through communication lines and terminals. In a few years, most of the catalog data in western languages will be available in the standard machine readable MARC format, and international standard book and journal numbers will be widely used. This will undoubtedly open new possibilities for systems linkages between jobbers and libraries.

Another very promising development is in vendor-supplied turnkey minicomputer systems capable of handling circulation, ordering, accounting, serials control, cataloging, and so on. CLSI, in Waltham, Massachusetts, appears to be the leader and is already offering book ordering, accounting, and circulation modules. It has made several installations, and has contracts for others, including a circulation system for the Harvard library. The entire system, hardware and software, is sold or leased to the library at prices which are advantageous when compared with the cost of developing, operating and maintaining in-house systems. The installation and maintenance of the turnkey system are the responsibility of the vendor who tailors a basic system to fit the needs of the particular library. The library with such a system is free from the vicissitudes of dealing with a university computing center or a library cooperative and retains a measure of independence and flexibility as the technology develops. CLSI will probably be joined by other vendors using a similar approach in the near future.

Information Design, in Palo Alto, is successfully supplying libraries with ready-made packages in computer output microfilm of the 400,000 records in the Library of Congress MARC data base. Because the data is sorted and appropriately arranged, catalog entries can be easily found and cards produced on an in-house reader-printer by a library staff with no technical expertise. Josten's, Inc. maintains the Library of Congress MARC data base on its computer and fills card orders from libraries on demand, thus making it unnecessary for small libraries to cope with any level of automation or technology.

Libraries can now buy the keyboarding of bibliographic data as well as computer-produced book catalogs and many other products and services.

As salaries and personnel benefits continue to rise, libraries will find it increasingly advantageous to buy these products and services rather than to attempt to produce them in-house. There is a future for imaginative vendors in the library automation field.

The third category covers *package systems* that have been developed by individual libraries or agencies and are available for and capable of being transferred to other environments. Several examples can be cited. Ohio State University's on-line circulation system is being considered for transfer to the SUNY system, and the University of Pennsylvania's batch processing-System 7 circulation system is being offered by IBM and Penn as a package to other libraries. The Hennipin County (Minn.) Library used the California Bibliographic Conversion System to convert its catalog entries, and then processed the entries with the New York Public Library's book catalog system to produce a catalog. These successful ventures suggest that systems transfer may finally be possible now that there are systems of good enough quality to transfer and technical people who are willing and able to do it. University libraries are beginning to overcome the "not invented here" syndrome, and systems transfer may begin to occur more frequently.

There is still a good deal of "do-it-yourself" library systems development going on in many libraries throughout the country where a library systems group composed of a few library or computing center systems people will undertake to design and implement a unique system for a particular library. Experience has shown that such systems are inordinately expensive to develop, maintain, and operate. Many libraries are quietly abandoning this approach in favor of joining cooperatives such as OCLC, or purchasing turnkey systems. The do-it-yourself approach will continue but will probably diminish in significance.

Major libraries such as the Library of Congress, New York Public Library, Stanford University, and the University of Chicago have done some of the best pioneering development work in the field, and, should not be classified among the do-it-yourself operations. Indeed, the University of Chicago systems group which is among the most experienced in the country is developing, with support from the Council on Library Resources and the National Endowment for the Humanities, one of the most conceptually advanced and flexible bibliographic systems in existence. Chicago's approach in its second generation system is to use a modular design which makes the on-line aspects, the data base management, and the applications programs all separate from each other and, to a large extent, independent of the peripheral hardware that is used. The system is being designed for an IBM 370/168 and has a minicomputer front end to act as data concentrator and communicator with the CPU. The basic data and file handling capability, which is called HERMES, has been designed so that it can be used by the development office, the hospital, and a number of other business applications in the university in addition to the library.

When the University of Chicago system is completed by the end of next year, it should be capable of being used either as a stand-alone system for a particular library or as the central system for a regional consortium. Thus it may be an example of a large and flexible transferrable system and as such would have an important influence on the future development of library networks.

Cooperative Automation for Libraries

by Russell Shank
Smithsonian Institution

Comments in this paper are based on the premise that cooperative development of automation, and the sharing of computer resources, is the most feasible approach to full utilization of the computer's potential by libraries in an academic and research environment. This was judged to be the case for federal libraries during a study of automation among them conducted for the Federal Library Committee by the System Development Corporation. The success of computer networking experiments, especially the development of packet switching, and new developments in telecommunications technology aimed specifically at the transmission of digital data also lend creditibility to this judgement.

There are three possible conditions under which academic libraries can successfully share computing resources. Libraries can join together in some administrative relationship, either to fund jointly the development of shared automated systems or to design a system that integrates software and computer services available from outside agencies. Alternatively, an individual library can build a custom computerized operation with facilities and services from outside sources. The consortium concept is the most promising of the three alternatives. Federal libraries in Washington, D.C. have joined together to form a network to access the data base for book cataloging at the Ohio College Library Center in Columbus, and hopefully to access data bases held elsewhere. These libraries may later turn to joint development of a cooperative system or to joint operation of consortium-owned computers. However, formation of the Washington, D.C. network does make some full-scale subsystems available quickly.

One must clearly differentiate between the sharing of computer resources through the diffusion of software packages for local use, and sharing through access to remote computers for whatever services their software produce. In the Federal library community network management assumed that individual internal operations could be adjusted to utilize systems run on remote computers faster than software packages could be modified to operate on local computers.

By relying on outside software and computerized information services for data handling, calculations, or information storage and retrieval, library management can avoid an expenditure for detailed design of performance systems, programming, and debugging. A library can introduce a sophisticated computer capability for one or another function relatively quickly, following the decision that such a function should be automated. Purchase of services, however, does not eliminate the need for systems analysis. Determination of the needs and purposes of automating, remains a large and vital task, as does the very difficult task of analyzing performance capabilities, design specifications and a host of administrative details relative to outside resources.

When one decides to make use of outside computing resources, one must rapidly alter internal operating procedures to accommodate that system. The long delay occasioned by the arduous tasks of inventing, testing, refining and documenting one's own computer system does not occur. The dimensions of the fall-back and recovery deliberations suddenly becomes larger and their delineation more urgent. One always has the spectre of the outside resource being taken off the market or changed in some significant way, leaving the sharing agencies with a void in vital internal operations. However, this fear should decrease as the number of computers to be shared in networks, and the functions they perform increase. If internal system performance specifications are clear, and if redundancy can be built into networks, switching to alternate resources may be quick and relatively painless.

In evaluating the utility of shared development of computerized operations, one should consider two factors: the number of functions to be automated and the number of agencies included in the automated system. The move to buy services and software for an integrated system from a number of outside vendors adds a third dimension to the model, thus increasing the sophistication of the analytical techniques required to solve problems. What look like constraints to viable systems from one side may be seen as advantages from the other. For example, as the number of libraries in an automated, on-line cataloging system increases, so do the problems of attaining agreement on standards for cataloging. However, as the number of libraries increases, so does the value of the data base for support of an interlibrary lending system. The two views can be expressed in a cost-benefit ratio, but there is no reliable and useful means for measuring the "value" of information services rendered in order to state the "benefit" side of the ratio.

The decision to achieve automation goals through shared computer resources solves some problems, but introduces others. In addition to the need for quick and accurate specification of goals, and the need for systems analysis, one of the most serious problems of working together is the negotiation of mutually acceptable goals and the establishment of priorities for the development of systems and subsystems. Compromises are difficult to achieve among strong libraries with different problems. The potential for system perturbations caused by changing membership in the consortium increase as the number of agencies involved increases. Agencies, both within the consortium and among the vendors of services, may be forced out of action by outside influences. Experiences among agencies that have attempted to build cooperative information storage and retrieval systems on a mix of data bases have raised serious questions, also, about the quality of data files.

In some cases there may well be no way to avoid the effort of innovating internally even if an institution wishes to share computer resources. This was the case with the Smithsonian Institution's development of an automated inventory system for museum collections, and a concommitant information storage and retrieval subsystem. The inventory problem was too heavy and pervasive, and no other museum seemed to be willing to tackle the tasks of introducing the computer as a tool for its solution. Since it could not go to the market place, the Smithsonian devoted several years to the task, developing its own system, training the staff, and investing in the necessary computer facilities. Now, the system is available for shared use.

Due in part to the telecommunications revolution in the United States, the impetus for shared use of computer resources is strong and will not disappear. Unit communication costs will go down as competition in the communication industry increases, making it more attractive than ever to create computer networks and information utilities for remote access, especially for libraries and museums. The combination of computers for processing and for communication channel management will enhance distributed computing concepts. Costs of developing and operating expensive central utilities can be shared, and those utilities can be used to drive considerably less expensive and limited processing by minicomputers at individual locations.

Chapter 19

Discipline Oriented Applications

Hierarchical Computing for Chemistry

by John Cornelius
University of California, San Diego

An understanding of the environment at the University of California, San Diego, is essential to an understanding of the approaches taken towards solving a number of fairly common data processing problems. In a Chemistry Department of 35 faculty members, approximately 20 are experimentalists who are either now acquiring computers, have one, or are asking for funds to acquire one. Each of the individual computers within the department can be justified in terms of the scientific mission of the various principle investigators. Sheer weight of numbers dictates that the department's commitment to automatic data processing in the laboratory and in subsequent data reduction is substantial.

In chemistry laboratories during the last five years, a number of turn-key systems have been integrated into instruments such as NMRs and spectrometers, and approximately six ad hoc laboratory systems have been built around either minicomputers or the departmental computer, an IBM 1800. Approximately eight additional minicomputers will be installed in the department by 1976, many of which will be turn-key systems included as part of an instrument that can be acquired at a price which is extremely competitive with alternative schemes such as utilizing the IBM 1800 or some existing minicomputer. Several others will be home-grown systems built around minicomputers and will be dedicated to specific tasks for which there are no existing turn-key systems.

Because the department exerts little or no control over an investigator who wishes to acquire a computer, there is no standard minicomputer in the Chemistry Department at San Diego unlike many computer centers, computer science departments, and chemistry departments elsewhere. Incidently, one does not need to view this diversity as being particularly bad. Although standardizing all minicomputer applications within a department or within a university campus is technically expedient, it is frequently extremely unpopular and presents many political barriers to fulfillment of the service function.

UCSD has a campus computer center which is in debt, a common situation in the university environment. The indebtedness of the campus computer center and the campus commitment to that computer center, which is required because of financial constraints within the university, produce some interesting side effects both politically and scientifically. Decisions regarding acquisition of a computer for use in categorical research by an individual investigator are often made on a political basis rather than on the merits of the problem or the merits of a particular machine that might be available.

Research dollars are becoming more and more difficult to get. In disciplines which can provide services and solutions such as cancer and air pollution, basic research is not being funded while more productive service type science is. Research proposals are being written for smaller amounts and the funding process is considerably more competitive than five years ago. This means that investigators have to take advantage of as many existing capabilities as possible in order to acquire funding.

The objectives set by one group of Chemistry faculty at UCSD for research into utilization of minicomputers are addressed directly to the problems of money and the computing capability that one can buy with it. A first objective is to promote innovative use of computers and advance the state of the art of computer use for chemistry research. Secondly, the group wants to enhance turn-key systems beyond the capabilities that are normally delivered by the manufacturer. For the fairly common price of $15,000 to $20,000, a turn-key system usually lacks adequate peripherals to do germain things like compiling, assembling, link editing and loading of object programs. In order to have the proper perhiperals for accomplishing these things, one must double, or quadruple in some cases, the cost of the minicomputer system which was acquired as part of a fairly inexpensive spectrometer in the first place. Third, a consistent level of systems support both in hardware and software should be provided for all the different kinds of minicomputers in the UCSD Chemistry Department. Fourth, the group would like to utilize the campus computer center whenver possible because it alleviates a lot of the pressures which could be put on an investigator to spend *all* research money for computing at the computer center. When there is no argument regarding what computer to use, it may as well be easy and natural for the investigator to use the campus center. On the other hand, use of the campus computer center should not be promoted in several areas since there are many instances where the campus center is simply not economical. One shouldn't promote the use of money in an unreasonable way since there is not enough of it. In areas where the computer center is unsympathetic or where expertise and programs clearly reside elsewhere, one should take advantage of that expertise. Parenthetically it should be noted that the bulk of support software that comes from minicomputers is written for IBM 360s. The UCSD campus computer center has a Burroughs 6700. Since there are a number of 360's on the ARPANET and the campus is on the ARPANET, Chemistry researchers utilize the remote IBM 360 computers by the ARPANET wherever it is reasonable to do so.

To provide maximum benefit possible for the investigators equipment and computing dollar, specific actions that are well within the state of the art have been taken to meet the objectives outlined above. No effort has been made to advance the state of the art in computer science in this project since the existing technology is adequate to provide the prerequisite hardware and software.

Any system which interconnects computers requires a great deal of interfacing effort. To make effective use of this effort, project staff at UCSD adopted a standard hardware interface unit which will provide relatively simple protocols for user and system hardware. The CAMAC data processing standard for computer interfaces has the advantage that it is not tied to any single manufacturer of main frames or peripheral devices. Essentially this standard is a peripheral device which serves to multiplex large numbers of peripheral units with diverse data transfer requirements into a single central processor. Since the protocols in CAMAC are simple, the UCSD chemistry computer facility can utilize personnel who are not trained in computer science or engineering. For the past several years, the project has been utilizing undergraduates, graduate students and post doctoral fellows who are not normally nimble with integrated circuits and digital design techniques. This is a decided advantage since the undergraduates, graduate students, and post doctoral fellows are the ones who define the problem area and often conceive the best solution to a particular problem. Utilizing this manpower in a relatively efficient manner has been extremely advantageous in both the application sense and in the system sense. CAMAC also reduces significantly the number of interfaces that have to be made to the numerous minicomputers in the Chemistry Department. At the moment the Department has 10 central processors representing 8 different manufacturers and 9 computer types. Using CAMAC allows faculty to plug things into any of these computers without building a special interface for each computer and each piece of hardware.

HIERARCHICAL COMPUTING

Utilizing CAMAC project staff has designed, and partially built, a computer system which provides bulk storage, utility services, peripheral devices, and other types of data processing support to minicomputers distributed throughout laboratories in the department. Presently two of the department's minicomputers are connected to a central IBM 1800 and utilize the peripheral devices and disk storage at that site. The two minicomputers, a Texas Instruments 980 and a Digital Equipment Corp. PDP-8 access the 50 megabyte disk storage on the IBM 1800 in real time and both computers operate under the manufacturer supplied disk operating system. Because the disk driver software provided by DEC and TI has been replaced with UCSD software that assumes the remote disk, the Department has saved approximately $30,000 by not buying disk for these two minicomputers but instead utilizing disk currently available in the Department.

Within the next three years with approximately 20 minicomputers planned to be connected to a central file system, the Department anticipates savings on the order of $200,000.

The central computing facility in the Chemistry Department also provides various utilities in the form of peripheral devices, disk management programs, and in some cases remote computing where it is not practical on the minicomputer. Many of the peripheral devices which are available at the central facility are interfaced through CAMAC permitting their utilization at the minicomputer site.

Hierarchical computing has actually been in existence much longer than the communications technology which has made possible the type of hierarchical computing that exists today. Transfer of information records between computer by manual methods still occurs, but is disappearing quickly. The advent of current communications technology permits higher speed data transfers between computers but the essential change has been in the boundary conditions. Today the turn around time between steps of a hierarchical computing system is several orders of magnitude faster than under manual methods. Nevertheless systems approaches are the same ones used under the early systems to transfer information records and programs, only the medium has changed.

REFERENCE

1. Research on this project is funded by the National Science Foundation Grant GJ36356, and the National Institutes of Health Grant RR00757.

Computer-Based Medical Records

by Elemer Gabrieli
SUNY at Buffalo

Although patient documentation is considered increasingly important, the medical records in the hospital chart are still informal notes, without explicit rules concerning content, format, and style. The ambulatory patient's office records are even more sketchy and expedient. This lack of standards is the first problem, since utility of computer-generated outputs is limited by the quality of the input records. For computerization, purposeful input is inoperative.

Computer-oriented structuring of the medical records followed an erratic path during the 1960's. Individual research projects developed their structured clinical records as a part of the ground work to launch the project. The design and composition of the data base reflected the opinion and bias of the clinicians involved. These projects focused on the technology, or on the output rather than on the input. It was often stated that the purpose of the model is to prove the feasibility of automated records. I would like to propose that the construction of good clinical records is still an unmet challenge.

In order to satisfy the data needs of the various users, the clinical input record must be *clinically oriented, rich in information, brief, computer-compatible, integrated* with all other records of the same patient, and the layout design must serve the interest of the data recording clinician. The latter is important, to *keep the clinician interested,* motivated. A return loop should be designed, with this psychological stimulus in mind. Further, the completion of the record should be simple, expedient, acceptable to the clinician.

RETRIEVAL – ARTIFICIAL MEMORY

The purpose of computerization is not only to retrieve a patient's record, a plain clerical function, but also to develop an external, artificial memory[1] which can support our overburdened memory. For such a man-machine relationship, we must be able to retrieve similar cases in form of summary statistics when "descriptors" are chosen by the user of the data system. Such an artificial memory must retrieve and tally single data elements from the vast data bank, crossing the individual patient record barriers. We may want to see the success/risk ratio of a certain drug in several diseases, in a certain age group, etc. For such a flexible information retrieval capability, we must store the clinical data individually, we must develop a network of organized data, to function as an artificial cognitive memory. The first criterion for such an ambitious undertaking is to *code* the input medical records. The machine must identify the semantic elements in the individual input record, and it must assign a purposeful code to each semantic entity. This calls for a national, uniformly used *medical vocabulary*. The urgent need for such a stable, formally controlled carefully designed medical vocabulary is quite apparent. Once such a vocabulary is completed, automated semantic coding of the input record will be the first step toward the creation of an artificial medical memory. The design of an immaginative, purposeful information processing blueprint will be the next large task of this decade. Retrieval-oriented processing of the submitted clinical record calls for separation of the input data into individual basic statements. Then, coding of semantic elements and linking the data to those in storage with corresponding meaning will be another programming step.

ARTIFICIAL MEMORY – POOLED EXPERIENCE

Retrieval of shared clinical experience is based on organized filing by semantic content of the input records. Retrieval can thus be unrestricted, since the semantic files represent *all* the terms used in our medical language. (There are about 35,000 basic medical terms currently in use.)

Medicine is the last profession still practiced from memory. All other information-dependent professions have changed to literature search prior to opinion formation. It is only natural that medicine has been unable to keep up with progress in research. The growing gap between available knowledge and its use at the bedside is a growing concern. Also, in medical education, it is increasingly difficult to adequately cover all traditional and recently developed areas in the curriculum. Educators must recognize that human memory has a limited capacity, and that the amount of facts a physician should know is already more than a human memory can store. We should no longer force medical students and physicians to memorize a vast amount of trivial data. Computer-assisted clinical practice would be substantially better if the machine could support our memory. Retrievable shared clinical experience should enable the clinician to review similar cases before selecting his diagnosis, and/or before organizing his therapeutic strategy. This man-machine memory combination will liberate the physician from his current limitation by his own memory, it will enable him to focus on the patient. Such a retrieval capability will also shift the criteria for selecting medical students. Instead of memory quality and capacity, human qualities and rational thinking should be the highest values when evaluating medical school applicants.

IMPLEMENTATION

The technology is available for implementation of a computer-based artificial medical memory. The computational linguistic aspects have been clarified, a computer-oriented medical vocabulary is now near to completion[2].

The Universities may play a major role in the implementation of a much-needed computer-assisted clinical system. The first task may well be to re-orient the teaching facility. An inherent prejudice must be overcome. Some of our leading academic clinicians may fear that machine-based memory would replace human thinking, individualization, clinical judgement. This fear is partially justified, since clinical function based only on external memory will be expected to be substantially different from our current practices. On the other hand, since the capacity of human memory limits clinical medicine to utilize currently available knowledge, we should reevaluate traditional values and upgrade bedside medicine. We should separate those functions which can be done better by a machine, from those which are the important human functions. This calls for honest soul searching and extensive reorientation.

An artificial clinical memory will potentially upgrade clinical decisions, but its efficacy will depend on the ability of the individual user to gather the clinical data comprehensively, accurately. Also, the interaction with the artificial clinical memory must be effective, productive. The dialogue comparing the clinical problem with the machine-stored cases must be logical, judicious, purposeful. The results should be limited only by the quality of the machine-memory, and the logical development of the rational clinical decision.

Universities should coordinate the large national task to develop the first artificial clinical memory. This requires a multi-disciplinary approach, probably a mult-university approach, since no university has more than a few experts on campus. The construction of such an electronic clinical memory calls for a joint team of clinicians and information scientists to design the blueprint for the hardware-software combination, and it requires a body of clinical experts, in every recognized field of clinical medicine to develop technology-oriented input records, to specify the output criteria, and to analyze and improve the semantic "isodata files", the foundation of the artificial memory. Unless academic medicine is willing to accept this vast challenge, by default, government and/or organized medicine will be forced to respond to the growing pressures. In the latter case, most probably the brain resources will still be recruited from the academic circles, but the team effort may be planned by medical organizations currently concerned with peer review. Such a development would erode the role of the academic system. Large influx of funds would strengthen the various "colleges", "academies" and "associations", rather than present university structures.

Another urgent academic task is to train physicians, nurses, medical record administrators, hospital administrators, information scientists, health care planner, some to staff clinical information centers, and the rest to use artificial clinical memory. Physicians with good clinical insight and with broad background in communication sciences and in information handling technology must be educated, as to a new type of clinical specialty, to implement, to maintain, and to continuously upgrade the computer-based artificial clinical memory.

Serious reevaluation of our entire medical school curriculum is also in order. Once we accepted the potential of automated information processing for upgrading clinical decisions, the selection of medical students and teaching faculty of the medical schools should be revised. Medical students should be trained to recognize similarities, and to be able to diagnose and treat using an artificial memory, i.e., to diagnose a disease without past experience with a certain drug. Such machine-supported thinking, with the electronic memory as an organic part of the medical decisions, requires carefully planned, gradual elimination of personal experience/bias. It also calls for subordinating our own recall to that of the machine with a much larger data volume, and with up-to-date, competently organized information base. The teaching faculty, and then in turn the student body must learn simple retrieval of pertinent information, and also how to form a clinical judgement based on accurate clinical data acquisition, followed by judicious comparison of the clinical data with similar experience offered by the machine. Rational decision making will no longer be limited by the clinician's memory and experience in corresponding cases but dependent only on the clinical qualities and logic of the decision maker.

Over the centuries, many new ideas developed in the academic mileau, and ideological-technical progress enriched our culture. Computerization of medicine seems to ignore this traditional pattern of evolution. Data centers mushroomed during the last decade, frequently apart from academic structures. These data centers often choose pragmatic goals, quite at variance with the conceptual values traditionally advocated by the academicians. Most of us considered the accounting-oriented hospital-based data centers as temporary arrangements, to last only until good medical data systems will evolve. We expected larger medical information systems to replace the current billing/inventory type systems, providing fiscal services as derivatives of the medical data system. In reality, there are already more than 3,000 data centers currently in operation. Most of these are completely separated from clinical medicine, teaching, and research. These data centers seem to become important factors in two aspects. One is the fact that some of these administrative data centers have accumulated a vast volume of medical data, and now these are the major sources of data for further planning of U.S. medicine. The other even more important aspect is that clinical medicine has lost control over the access to such data banks. Patient privacy is now in jeopardy. It is proposed that with adequate funding, a powerful artificial clinical memory could be implemented within two or three years. However, before we agree to automate sensitive clinical information, we must assure our patients that their right to privacy will be honored. If patients in the consultation room hesitate to reveal their complaints, or fears, if our patients begin to worry that their private statements may "create a record", if an uncontrolled data system can become a potential source of embarrassment or humiliation, our entire social structure will change. Academic medicine should play a critical role in this sudden crisis of confidentiality. Computers penetrated our health care system, peer review requirements will further encourage automation of medical data. Simultaneously, the growing role of government in paying for health care, and the justified need for data to control cost and quality, are all added pressures to augment the scope of automated medical data. In the name of economy, or political expediency, our most cherished western cultural values and civil rights may be lost. Academic medicine must become the champion of our cultural and political heritage. Medical privacy, patient confidentiality, are not in conflict with data automation, but unless academic medicine develops the ethical guidelines, the legal framework, the data security standards, medical data will be increasingly computerized and used for many justifiably purposes, ignoring the privacy of the patients.

Universities are now in a peculiar situation. Computerization within the health industry has developed unplanned, uncoordinated. Our academic posture is still traditional yet computer technology is already a part of medicine. To derive the potential benefits of the technology, and to control the inherent risks, bold national planning is necessary. It is proposed that academic medicine should sit at the planning table, and become a visible leader in progress, and the champion of our cultural heritage.

REFERENCES

1. Gabrieli, E. R., Conceptual Aspects of a Computer-compatible, Stable, and Controlled Medical Vocabulary; *J. Clin. Comp.*, Vol. 3, 1973, p. 2-23.

2. Transactions of the Conference on Computational Linguistics, *J. Clin. Comp.* Vol 3, Special issue, 1973.

User Issues in Medical Record Automation

by Robert M. Strong
Lincoln Laboratory, M.I.T.

A few practical questions have shaped the design of an automated medical record system being implemented in a small, less than 10,000 patients, primary care setting which is part of the Outpatient Department of Beth Israel Hospital in Boston, Massachusetts.[1] Work is being done on a much more limited scale than that which Dr. Gabrieli has suggested, and should provide a contrast of the practical with the ideal. To some extent the project has met his criteria and laid the groundwork for the future.

The implementation is being carried out using the computer facilities of a record system developed by Drs. Grossman and Barnett at Massachusetts General Hospital for the Harvard Community Health Plan.[2] The decision to take this route reflects staff feeling that sufficient work has now been done on the computer science aspects of such systems and that further work will have at best a second-order effect on cost and responsiveness, at least for systems handling less than 30 to 40,000 patients.

The first-order problems, from the point of view of those users whose primary concern is care delivery, are cost, availability of the data at the time of a patient visit, and the simplicity of the data entry mechanisms. Why the concern at this level? It can well be argued that the most important reasons for automation involve use of the data for non-clinical purposes such as epidemiology, studies of utilization patterns for academic or management purposes, audit of provider performance for quality control, management of costs in prepaid settings by control of provider performance, and so on. Whether one's interest is in the academic vein, or in a specific management problem, be it management of an OPD, an HMO or the Medicare payment mechanism, the data comes from the patient-physician encounter and the events which stem from it. Collecting the data depends on the cooperation of the clinicians on the firing line. It is a shock to find that in any real setting that is precisely how the clinicians view their position; as being on the firing line where the delivery of care to patients is the *only* goal. This seems to be true even in heavily academic settings like the Beth Israel Hospital.

At Beth Israel, fundamental reasons for automation are non-clinical, and there is little chance that there will be an immediate improvement in clinical care as a result of automation. In the long run, better management must result in better and cheaper care if the management values are selected properly. The availability of this data must result in an improvement in understanding of disease processes, and sooner or later the barriers to record linkage and patient-portable records must fall. However, the clinician cannot now justify any burden to the patient care process on this basis. He or she must look for immediate benefits in return for present burdens.

COSTS

The economic impact of an automated record must be kept small. Typical current paper record costs range from 50¢ per visit to $2.00 per visit, depending on whether the setting is a private office or the out-patient department of a teaching hospital. At the Beth Israel, costs for keeping medical records have historically averaged $1.72 per visit. Typical estimated equivalent costs for existing automated systems suitable for small populations run from $3.00 to $5.00 per visit. Is there any way one can justify paying *more* for an automated system than one can save on the replaced paper system? I think there is. Until recently, staff on the Beth Israel project were trying to convince skeptical administrators that the data made available through automation really would allow them to save money. Numerous arguments were used depending on the incentives operative in the particular setting. However, now one can make those arguments on the basis of real data that money can be saved. Drs. Schroeder, Kenders, et al[3] offer data suggesting that physician performance can be modified in significant ways by peer pressure. Their data suggest that for their selected patient population, physician behavior could be modified to the tune of $15.00/yr per patient averaged over drug orders and laboratory procedures.

If this number is correct and repeatable in other settings, even allowing for a factor of 2 error, it implies that one can pay for the automated record instantly! In spite of the unusual selection procedures of Dr. Schroeder's experiment, data indicate that results can be achieved in many other settings, and for much different patient populations. In prepaid settings, normal economic incentives are more likely to be operative, and the drives to manage costs are likely to be stronger. There is yet no proof that physician performance can be altered on the basis of such goals alone. However, the Harvard Community Health Plan has been engaged in an attempt to use data derived from their record system to alter physician habits with a goal of improved quality and seems to have had considerable success. Perhaps one need not force the added costs of automation to zero, but need only to make the cost reasonable. The cost target at Beth Israel Hospital is an increase of no more than 30%.

THE MARK-SENSE TECHNIQUE

There are some techniques which can be used to control costs in automated record systems. A quick analysis of existing systems shows that the largest segment of the system cost is associated with collection of the data. This effect is always present in both the computer budget and the personnel budget. Some researchers, anticipating a drop in the cost of computer hardware, have proposed an intimate relationship between the physician and the computer with a CRT terminal in every office. One effect of such an arrangement is to obscure the cost of the personnel involved, the physician. A second effect is to interject a disruptive influence into the physician-patient relationship.

In the Beth Israel project, an attempt has been made to minimize both the computer costs and the personnel costs by collecting much of the data using mark-sense forms. The physician deals entirely with a *paper* record. While almost all of the duplication of data recording, and most of the clerk intermediaries have been eliminated, the computer costs have been reduced. The intent is to collect all possible data via this mechanism. Currently the design includes: collection of problem list entries and changes (it is a problem-oriented record); laboratory orders; medication orders; encounter data; some social data; and the results of protocol directed encounters with various non-physician providers. The ASC group is also engaged in the development of such protocols.

There are, of course, serious drawbacks to the mark-sense method. The most important limitation is that any list must be finite in length. Since the set of comments which a physician might desire to enter into the record is semi-infinite, one must draw the line at some point and accept as a given that there will be some free text entry by clerk. Project staff believe, however, that in those areas chosen for mark-sense entry more than 90% of the entries made can be handled by the mark-sense method. Use of the mark-sense method entails the development of different forms for different provider classes and specialties, but this burden is not as great as that imposed if more data were entered by the clerk.

The forms design problem turns out to be a non-trivial one. Lists must be limited lest they become ungainly and difficult to use. Remember that acceptance, or lack of it, by the providers ultimately determines the amount and quality of the data collected. Lists which are too short soon cease to serve the provider's needs and the system will be subverted in unimaginable ways. Data entry costs are essentially proportional to the data which must be entered by clerk.

One critical problem has not yet been resolved in the Beth Israel project. What can be done about the progress note, the catchall for other data which the physician desires to enter? Basically, we don't know yet. It seems clear that most of the immediate goals of automation can be achieved on the basis of the data already collected. It is unlikely that physicians will provide much of this data in coded or codable form in the near future, although one can attempt to collect signs and symptoms via mark-sense techniques. There is no hope of being able to do an automated analysis of free text notes. Part of the difficulty with progress notes stems from the physician's view of his role in primary care settings. Unlike the hospital based practitioner or the specialist, providers in the primary care setting see themselves as caring for the patient, not for his diseases. Consequently, data relating to the patient, and to his or her state of mind and general well being are more important to the physician that the state of the patient's diabetes or hypertension. One is often criticized for attempting to "take the *life* out of the record," a concern for the life of the record is not important in any of the non-clinical uses which have been contemplated by the Beth Israel project staff, and perhaps would not be as important in a hospital or specialty practice setting. However, it is important to the physician in primary care.

The primary non-clinical use of the progress note which has been suggested is audit of physician performance by the tracer disease method. For other reasons, the Beth Israel record system would require that such audit, as well as many other longitudinal studies, be done by hand. This is not viewed as a critical constraint. From the point of view of the clinical uses of the record, the primary reason for being concerned with the progress note is cost. In simplistic terms, the presence of the progress notes on the printout is intended to save money by eliminating some of the activity of the paper record, or eliminating the paper record entirely. Until better methods of computer storage are available, capable of handling non-verbal material such as ECG's, X-rays, and occasional photographs, there will always be a need for some kind of paper file. Project staff have chosen to retain the paper record and to use it as the repository of all progress notes. It is to be pulled only for the patient visit and to file data which cannot be entered directly into the automated record. The physician's progress note is written directly into the paper record.

An examination of the activity of the paper records of active outpatients at the Beth Israel Hospital revealed that 70 per cent of the activity in the record room could be eliminated by two simple mechanisms: the use of the computer summary printout for review of returning laboratory data rather than pulling the paper records; and the use of an automated reporting system by which the hospital's laboratory information system computers report results directly to the automated record. The summary printout includes the problem list, lab data, medications and visit history. If the hospital computer reports laboratory results directly to the medical record system, it is not necessary to employ a clerk to return paper reports to a paper record.

THE SUMMARY RECORD

Hopefully the limited summary record will be sufficient source data for a large fraction of the unscheduled visits made by the Beth Israel patient population. If this procedure is satisfactory, the center can save the cost of the associated paper record activity.

There is one more important characteristic of the Beth Israel system. Data on the Summary is viewed as having a limited useful lifetime. After a suitable period, now set at 12 months, it will no longer appear on the Summary unless "current" in some sense. Current data includes: medications ordered and not discontinued; laboratory data which is the most recent example of its kind; and problem list transactions which are always current. At the time the data is expunged from the computer file, it is placed on tape and the paper record is brought completely up to date. The expunging procedure is designed primarily to limit the length of the Summary and make the important data more easily available to the physician. It also has the effect of conserving disk storage space, limiting the maximum access time for an unscheduled record request and conserving computer resources at both entry and print time. In addition there are other, less desirable, results of the policy. On-line retrospective data search will be limited to one year's data. Tracer type audit will be very difficult to do. One can expect some calls for the paper record in unscheduled situations if only because of physician insecurities. Sign and symptom data, when available to all, will probably not be extremely accurate. However, these problems are not particularly burdensom given project goals. The system will contain a prospective data tabulating provision and this, in conjunction with the 12 months of current data should be sufficient for most anticipated needs.

Data compression is not a problem primarily because the Beth Israel system is isolated. Techniques for the compression of data in known formats are well known, and since disk storage costs are dropping rapidly the value of further work is doubtful. There is need for work in the area of coding diagnoses, signs, and symptoms, so that data retrieval and automatic analysis can be carried out. This is a perennial problem which is one not of assigning code symbols to terms, but one of standardizing terminology. Medical language is full of ambiguities and duplications. Furthermore it tends to be descriptive rather than nominative, particularly when the problem-oriented record is is use. This is an enormous problem. Anyone can establish a set of terms for his or her own use and assign a suitable code for internal use. However, a multiplicity of terms does not permit interchange of information or enhance the comparability of data.

The Beth Israel project is not working in this area although solutions are needed. There are two efforts worthy of note: the SNOMed code which is a good compromise between classification and description; and the CLIP system in development by Drs. Simon and Leeming at the Beth Israel Hospital. Dr. Gabrieli has described SNOMed earlier in this Chapter. Its principle use will be in the coding of diagnostic terms and problems in a problem oriented record. It will form the basis for disease related retrieval mechanisms, and hopefully will become a universal standard for that purpose. CLIP is a purely descriptive language with a limited vocabulary designed for the characterization of radiographs. The primary advantage of CLIP is that it is designed to permit easy conversion of coded statements into more pleasing English text for printouts. CLIP appears easily extendable to coding of other consultant reports, such as EKG's, and may be the structure needed for coding progress notes as well.

Work at the Beth Israel Hospital is just now getting off the ground. Hopefully, the project will succeed in developing a system which works at the Beth Israel, and is directly applicable to other small settings as well.

REFERENCES

1. This report was performed under Contract HSM 110-73-335 with the Bureau of Health Services Research and Evaluation, Health Resources Administration, Public Health Service, Department of Health, Education and Welfare.

2. Grossman, et al. "An Automated Medical Record System," *Journal of the American Medical Association*, June 18, 1973, Vol. 224, No. 12.

3. S. A. Schroeder, K. Kenders, J. K. Cooper, T. E. Piemme, "Use of Laboratory Tests and Pharmaceuticals – Variation Among Physicians and Effect of Cost Audit on Subsequent Use, *Journal of the American Medical Association*, August 20, 1973, Vol. 225, No. 8, p. 969-973.

Facilities for Data Manipulation in the Humanities

by Michael E. Lesk
Bell Laboratories

The variation in computing services available to humanities researchers is very broad, and not entirely related to their ability to pay. Cooperation among humanists should result in a wider awareness of useful computer facilities, techniques and services so that all could perform their research with computer programs of the power now available at only a few sites.

This discussion is restricted to written data, ignoring the interesting but difficult problems presented by digitized speech and music, machine-readable forms of two-dimensional pictures, and descriptions of physical objects such as museum holdings. Fewer general techniques exist for data other than written language, and even the few tend to be both more voluminous and more expensive, and thus of interest to a greatly restricted set of researchers. Writing is more compact and more easily processed. A typical user of time-sharing system at Bell Laboratories uses about 100,000 bytes of on-line storage for his programs and data. This represents only trivial amounts of humanistic data, perhaps 20 pages of printed text or 10 seconds of digitized speech, and only a tenth of one high quality picture. Even writing is quite bulky by normal computer standards, and other forms of data humanistic are frequently impossible to store on-line at all.

The wide differences in the facilities available to humanists have unfortunate effects in the transferability of work among workers and locations. Often, for example, the most interesting data are available from projects that have been operating for a substantial time; this may mean that they are using out-of-date formats and equipment. All too few installations can easily convert data from older computers into modern formats.

THE IMPORTANCE OF IDENTIFICATION

What sorts of data should be collected? There is plenty of machine-readable text. Various abstracting services, such as Chemical Abstracts, U. S. Government Research and Development Reports, and Computer and Control Abstracts, are distributing many thousands of abstracts from different subject areas. The production of books and magazines by computer composition has as its byproduct machine readable forms of many items. However, samples of raw text without adequate identification are of little interest. Consider the problem of identifying in a set of abstracts in English and German those which are translations of another abstract. Needless to say, this was done by hand. Given a set of paragraphs from the galley proofs of a computer-typeset magazine, can one assemble them into the original articles? It is depressing to see how often a potentially valuable set of machine-readable data is made useless for future research by the elimination of identifying or classifying information which was simply not of interest to the collector. In particular, for information retrieval applications it is important to have such information as the bibliographic references in a machine-readable article, any known questions referring to an article, and assessments of relevance where available. Reconstructing these linkages later can be painful. However, when a collection is properly identified and formatted, it can often be used by later researchers. Repeated use of the same material offers great opportunities both for saving work in transcription to machine-readable form and in the accumulation of data for further studies.

FACILITIES NEEDED

What facilities should be made available to assist in data processing, once the data are collected? First of all, on-line access with a time-sharing system is a great convenience. Those who only process magnetic tapes serially are in the position of a researcher who is never allowed to skim books in a library, but must read each one he takes off the shelf from cover to cover before he may look at the next book. Admittedly, storage requirements are high, but costs are rapidly coming down. Minicomputer systems should be examined by researchers: they are typically closer to the state of the art than large systems, permitting lower costs, higher reliability, and bigger capacities. With minicomputers overhead, both of the computer sort and of the management sort, is lower. In exchange, one loses the services supported by overhead, which one may want. The common ways of conserving space, such as organizing files without extraneous blanks or fill characters, or using compression techniques, normally yield only a two- or three-fold reduction in size. In practice this is often not enough to make an important difference, and the best solution is an efficient backup storage scheme providing for fast and automatic transfer of archived material to the working store.

The next important step, once the data are available in machine readable form, is to provide a good context editor software package to go through it. All too often, a canned editor is restricted to operation on line-numbered programs; for humanities use, an editor must have context search and substitution capabilities. Also a flexible editor can be used for many format changes that would otherwise require special text processing programs. To convert, for example, between a representation of upper and lower case in terms of case shift characters and in terms of an expanded character set should be a trivial operation in a good editor. Is is necessary to line-number files on tape or disk? One doesn't have to worry about getting them out of order when one drops them, and a good editor will find lines even if unnumbered. Good utilities are another requirement; this should include record-keeping facilities to keep track of past generations of data, and basic conversions between data media and data representations. Usually, the computer manufacturer supplies these, since business programmers need such facilities too, but there is at least one large computer system on which the manufacturer's utility programs could not concatenate two files into one. Among the obvious programs that may be wanted for humanities purposes are: word, letter, and syllable counters; sorts and concordances; and various kinds of search routines. In fact, a complete document retrieval package working by word matching may be of use in selecting passages for browsing if it is available.

A major problem facing humanities researchers is the preparation of high quality output. Many text formatting software packages are now available. Combining these with upper/lower case printers or terminals is almost a necessity. Many cheap terminals with good quality printing are available. For more exotic requirements, various devices such as CRT recorders, microfilm plotters, and dot matrix printers with programmable character sets can be used. These can be obtained at reasonable costs if high speed is not required. A particularly attractive alternative is the phototypesetter. Small phototypesetting machines can be obtained for less than the cost of a line printer and produce publication-quality output. With many such machines, one can order custom designed fonts of characters taken from individual artwork. For a few thousand dollars extra, one can have a machine which can print almost any alphabetic or syllabic language. With some typesetters based on CRT character generation, the character design can even be reduced to a software problem. Even if one's computing center does not offer a phototypesetting service, outside firms can often be hired for small runs.

To encourage and simplify the job of browsing through data, normally quite inconvenient on computers, a basic retrieval program may be of use. Even a simple selection program that finds occurrences of individual words or letter patterns can be of great help in locating relevant material (as well as a killer at Scrabble). More sophisticated retrieval programs can identify word co-occurrences and other structures of particular interest in documents. In these ways, the amount of time spent looking for what to study can be reduced, and the amount of productive time spent studying it can be increased.

Steps Toward Making Literary Texts Available

by Stephen V. F. Waite
Dartmouth College

Among the first recorded uses of computers to study literary texts are the efforts of John W. Ellison in 1947 analyzing the manuscript tradition of the New Testament and Roberto Busa's work on the campus of Saint Thomas Acquinas, begun in 1949. After a decade or more of slow growth, the past dozen years have witnessed an efflorescence of work rapidly expanding in both the number of those involved and in the varied nature of their interests.

Representative of the nature of the work being undertaken are projects as disparate as concordances and authorship studies. Since the establishment of the Cornell series of concordances in the late 1950's[1], computer-produced concordances have become accepted even by the people otherwise opposed to the intrusion of Vulcan into the houses of the Muses. While the first results were most often directly reproduced from printout, in uppercase only, since 1968 it has been possible to photocompose elegant books indistinguishable from those typeset normally, except for their accuracy. Lexical work for large scale projects in languages from French to ancient Greek is frequently predicated on the preparation of large bodies of text in machine-readable form. Grammatical studies, metrical work, and stylometric undertakings in general can be far more accurate and thorough than would have been possible without computer aid; since the amount of human effort expended on a particular project may well be a constant, the end product may not come sooner for the computer aid but rather be more comprehensive. Authorship studies, particularly those of Mosteller and Wallace on the Federalist Papers, have attracted wide attention.

This quick survey of fields of application is not intended in any way to be exhaustive but rather to point out one thing which they all have in common. All spring from the text, which must in one way or another be in computer-usable format before further work can proceed. Unfortunately, many humanities scholars are more facile in argument than in typing, and richer in imagination than in funds to hire assistants. As a result, many projects have foundered precisely upon the rocky problem of getting an accurate text; yet for other authors, as many as three projects have been engaged in the identical task of copying a text into computer-readable format.

For these reasons, there has been a growing undercurrent of interest directed towards keeping worthwhile projects from the aforementioned shoals by making those texts which presently exist in machine-readable form as widely accessible as possible. There will always be those who prefer to keep the results of their efforts to themselves, but experience to date has shown that such people are definitely in the minority. Various methods towards accomplishing the goals can be suggested, but it is unclear which of them will best avoid the multi-formed Scylla of intensely competing undertakings while skirting at a safe distance the Charybdis of endless planning.

The minimum that can be done is publicizing as much as possible the availability of those texts which are presently in machine-readable form. To a considerable extent, *Computers and the Humanities*[2] is performing this project for literary studies as a whole, while specialized material can be found in *Computers and Medieval Data Processing*[3] for medieval material, *Arithmoi*[4] for Biblical studies, and *Calculi*[5] for Greek and Latin. Although no newsletter is being published, information on early German is being collected by Rudolf Hirschmann, Department of German, University of Southern California, Los Angeles, California 90024.

Yet there is a vast gulf between knowing that a text exists and being able to obtain it. Faculty members change institutions, lose interest, and go on sabbatical; the last is perhaps most difficult to overcome, for a person who is in Europe is effectively isolated from tapes in his office in Baltimore. Further, literate individuals may justifiably have absolutely no knowledge of how a computer reads tape or cards, and the differences between the expectations of different installations can prove an extremely effective barrier. The next level of possible operation is that of having a center, however established, which undertakes to ascertain the formats required by various kinds of machines and acts as an intermediary for requests so that the material would be prepared appropriately for transmission from one installation to another. This approach, essentially that of acting as a broker, has never been tried.

The next step up in the scale seems to be that of having the texts themselves in some central location, whence they can be distributed upon request. This approach was adopted in the summer of 1969 by the American Philological Association, the national association of classical scholars, and much of what is said here is based on prejudices gained as a result of supervising the American Philological Association's Repository of Greek and Latin Texts in Machine-Readable Form. The Directors of the Association at the outset laid down certain guidelines: the texts are distributed at cost to anybody requesting them; there is no guarantee of their accuracy; and the work of the person who originally prepared the material in machine-readable form must be acknowledged appropriately in every subsequent use made of it. The Association has an Advisory Committee for Computer Activities, which has been helpful in overseeing the growth of the collection during the past four years until it now totals some 200,000 lines of text. Such Greek works as all of Homer and the New Testament, and Latin writings ranging from Cato the Elder to Boethius are represented. Tapes have been distributed widely both in this country and in Europe. In 1973 alone, there have been thirteen requests filled from five countries. Often many texts are included in a single request, so that it is safe to say that man-years' of effort have been saved. The costs for a tape are usually quite modest; a tape of the *Iliad* would normally be available for less than $20.00. It must be pointed out here that Dartmouth College has been most generous in its policy towards allowing computing time and that the project received a grant of nearly $10,000 from the National Endowment for the Humanities for the fifteen-month period ending in June 1972. Without both of these sources of support, work would have been far slower.

Publicity has been given to the collection in various publications and meetings, and a list of the holdings is available upon request. One additional possibility is that of including the listings of the holdings in the card catalog of the college library; this idea was first suggested by the college librarian, and it raises serious questions about the nature of accessibility and about the problem of depending on the knowledge that a particular text is today on one tape while tomorrow a corrected version may be on another tape. The various ramifications are going to be explored in the coming year.

One of the big problems that has been faced by the collection is that of standardizing the texts, especially those in Greek, which must be transliterated into a Roman font for printouts on most kinds of computer equipment. Different conventions have been followed by different individuals, often with eminent justification; the precise character used to represent a particular letter does not matter as long as one standard can easily be transformed into another. More important is the kind of information to be included; for instance, the British policy is normally to omit accents, while the American standard is to include them. For this reason, a 15-page set of guidelines for keypunchers was prepared this summer in conjunction with the punching of the entire corpus of Greek literature being directed by Theodore F. Brunner at the Thesaurus Linguae Graecae, University of California, Irving, California 92664; when completed, the results are to form the basis for major lexicographical work. These guidelines, while they cannot be all-inclusive and do not follow exactly the representations used by the American Philological Association, are compatible with them. Experience points up the desirability of having a uniform text format so that different works can be handled readily. While it is far too soon to begin even to think about standardizing the approaches being used to study the texts, standardization of the texts themselves seems inevitable, and, in Europe, cooperation on these lines is proceeding between centers in France and Italy for modern languages.

Inevitably also, errors are detected in texts in the Repository, and some recipients have been most conscientious about reporting them; one, Henrietta Warwick, of Minneapolis, Minnesota has even gone to the effort of obtaining the alterations which are to appear in the forthcoming third printing of R. A. B. Mynors' Oxford Classical Text of Vergil's *Aeneid.* These corrections, in good conscience, must be put into the versions being sent out, and careful verification must be carried out to make certain that the changes are indeed corrections. Certain kinds of errors, such as illegal collocations of characters, can be detected by computers, and relatively crude scanning programs based on simple metrical rules can show up mistakes in poetic texts. All of these proofreadings, editings, and corrections can take time to run, verify, and enter. The American Philological Association has been fortunate in having assistants, one of whom came on a volunteer basis, to help in this work from time to time.

The aim of the collection is sending out texts so that scholars and students at various institutions can exercise their own ingenuity in approaching their own problems. In effect, it is a network, one which depends largely on the U.S. Mail; as a sidelight, unlike most computer tapes, tapes of literary texts can be sent book-rate. Given the demands even of publish-and-perish tenure decisions, this relatively slow pace remains practical; indeed, for European contacts, it seems the only one practical. Yet literary scholars can look forward to piggybacking on a potentially existing national network designed for other purposes to allow yet quicker access to these texts.

Perhaps less desirable, such a network will also allow access to the programs which presently exist to manipulate the texts. Already, there have been requests from such philosophically and geographically disparate places as Madrid and Wellesley for concordances prepared on demand at Dartmouth. A network with the ability to transport easily 10,000 lines and more of results would perhaps make such output more readily accessible and cheaper than its present rate of about one-half cent for each word in the text concorded. On the other hand, it might lead to increasing dependence on canned programs with the resulting channeling of scholars away from fresh approaches. While the time for a national center for working with texts in the humanities still seems far in the future, Dartmouth has begun tentatively collecting texts in modern languages into a data bank with the acronym LIBRI, Literary Information Base for Research and Instruction. This collection will supplement the one for Greek and Latin.

Many technical problems clearly have been glossed over in this quick survey of possible coordinating centers. Equally important are considerations of copyright; here, there is hope that within a few months a major publisher may come forward with a release which would enable work for scholarly purposes and serve as a model for others to follow. Also untouched are other questions: how to establish centers to cover texts in modern languages; how to gain the support of professional associations for them; moral if not financial; and whether there should be one such center or several perhaps oriented towards specific languages or periods. Much remains to be done before there can be an advance beyond a user services type of network to a facilitating or transmission network, to use the terms presented at the 1973 Spring EDUCOM Conference[6]. Nonetheless, the keel laid now can be the basis of a shipshape product able to navigate safely the perils and problems described.

REFERENCES

1. Stephen M. Parrish, *A Concordance to the Poems of Matthew Arnold.* (Cornell University Press, Ithaca, New York) 1959.

2. *Computers and the Humanities,* edited by Joseph Raben, (Queens College, Flushing, New York 11367)

3. *Computers and Medieval Data Processing,* edited by Jean Gagne, (Institut d'Etudes Medievales, Universite de Montreal, C.P. 6128, Montreal, Quebec, Canada)

4. *Arithmoi,* edited by Richard Whitaker, (Central College, Pella, Iowa 50219)

5. *Calculi,* edited by Stephen V. F. Waite, (c/o Department of Classics, Dartmouth College, Hanover, New Hampshire 03755)

6. Carolyn Landis, "Planning for National Networking: A Report on the Spring 1973 Conference," *EDUCOM Bulletin,* Vol. 8, No. 2 (Summer 1972), pp. 2-8, particularly p. 3.

PART VI

A BROADER VIEW

315/316

Chapter 20

Data Transmission Services

Regulation of Computer Communications

by Robert P. Bigelow, Esquire

Government regulation of computer communications is a factor with which educational administrators must reckon in 1973 and beyond. Although federal agencies (executive, legislative, judicial and independent) are the most prominent promulgators of rules and regulations governing communications, states and industry associations are also involved. This paper surveys the significant factors and actions in the regulation of communications.

Regulation of communications in the United States is based on statute, the major one being the Communications Act of 1934. Other statutes that may be important include the Communications Satellite Act of 1962, the Automatic Data Processing Act of 1965 (known as the Brooks Bill) and the Antitrust Acts (Sherman, Clayton and Robinson-Patman).

THE FEDERAL GOVERNMENT

The Federal Communications Commission (FCC) is the major *independent agency* concerned with communications regulations. The FCC is charged with the "control of interstate and foreign communication by wire or radio" and the transmission of energy by radio. A common carrier (a company which serves the public generally) must furnish communications services at reasonable prices in response to reasonable requests. Within the FCC, the Common Carrier Bureau oversees the activities of AT&T, GT&E, approximately five hundred independent telephone companies, and the data transmission companies offering specialized services such as packet communications and microwave.

Other administrative agencies also have regulatory responsibilities. For example, the General Services Administration supervises the federal government's purchases of computers and related services; the Federal Trade Commission has issued orders regulating programming schools; and the Securities and Exchange Commission oversees certain financial aspects of public companies that provide computer or communications services.

In the *Executive branch* several agencies and offices play a variety of roles in the regulation of digital data communications. The Office of Telecommunications Policy, part of the President's staff, coordinates government policy regarding communications. In addition to his function as the principal advisor to the President on domestic and foreign telecommunications, the Director of OTP supervises a staff which provides liaison to the FCC and congressional offices, coordinates the operations of the federal communications system, and administers the national communications system (the federal system linked with other systems) in a national emergency.

Within the Department of Commerce, three groups are concerned with telecommunications. The Office of Telecommunications provides statistical support to OTP. The Institute for Computer Science and Technology in the National Bureau of Standards conducts special studies for various branches and agencies of the government. (This office may be most familiar to EDUCOM members because the director, Ruth Davis is an EDUCOM trustee.) And the Patent Office attempts to safeguard proprietary rights in computer software.

Also in the Executive Branch one finds other organizations that regulate various aspects of computer communications, such as the Defense Communications Agency of the Department of Defense, the State Department (duties under the Cable Landing Act), the Antitrust Division of the Department of Justice which enforces the antimonopoly laws, and the Treasury Department which promulgates IRS policies on equipment depreciation.

Congress itself controls one agency which has some responsibility for communications: the General Accounting Office headed by the Comptroller General who renders opinions on the legalities involved in the federal procurement of supplies and equipment, including communications and computers. Institutional and corporate lawyers may find helpful precedents in his opinions on the validity of corporate contracts. Congress, of course, passed the FCC Act and more recently Senator Hart has proposed a far ranging investigation of the concentration of power in the computer and communications industries.

The *Judiciary* reviews, on appeal, rulings by the FCC and other agencies and the Executive Branch, and decides cases brought by government or private citizens (for example antitrust). An important judicial precedent in computer communications is the Consent Decree entered in 1956 in the case of United States vs. Western Electric, where it was agreed that the Bell system would not engage in any business "other than the furnishing of common carrier communications services," with certain exceptions, 1 CLSR 24,30.

OTHER FORCES

Several industry associations also affect the regulation of computers and communications by their suggestions and, sometimes, their lobbying activities for different industry suppliers and users. The International Communications Association (ICA) represents large companies that are users of voice communications, and the timesharing section of ADAPSO represent users of data communications. The United States Independent Telephone Association (USITA) represents the non-Bell non-GTE companies. The North American Telephone Association (NATA) and the Independent Data Communications Manufacturers Association (IDCMA) represent interconnect equipment suppliers. The Computer and Business Equipment Manufacturers Association (CBEMA) represent the larger manufacturers, while the Computer Industry Association represents the small independent computer manufacturers.

In September and October of 1973 state public utilities commissions have become involved in regulating computer communications. In Nebraska, the Attorney General has taken the position that the owner of a private system must qualify as a common carrier. In North Carolina a similar proposal is under consideration before the state Public Utilities Commission.

REGULATION IN TRANSITION

Regulations governing the provision and use of communications facilities are constantly being revised by federal and state commissions. To understand the process of regulation, it is helpful to recognize the difference between a tariff (rules and rate schedule) which is submitted by a communications company and the tariff which carries the force of regulation. (The procedures at federal and state level are usually similar.) Tariffs filed by public utility companies are nothing more than proposed schedules of rates and regulation. They become effective after a predetermined time unless suspended by the commission. In most cases the commission itself does not consider the tariff (there may be a staff review), and its validity is not challenged. In this case the tariff becomes effective, but may still be challenged years later. It is only after a hearing and an agency decision that a particular tariff provision can be considered as "approved" by the agency.

Changes in regulations in recent years have focused on three areas: interconnection, transmission, and message switching. The Carterfone decision, 13 FCC 2d 420, 1 CLSR 1019, is perhaps the best known case involving the interconnection of non-Bell equipment with the Bell communications system. In Carterfone the FCC ruled that the Bell tariff absolutely prohibiting the interconnection of non-Bell equipment to the Bell communications system was illegal and always had been, but also said that non-Bell equipment could not be attached in a way that would harm the network. Notwithstanding continuing studies and committees to achieve interconnect standards, Bell and General Telephone have carried on war against the interconnection concept. Although individual FCC decisions like Carterfone have encouraged those interested in promoting interconnection, recent state activities, particularly in Nebraska and North Carolina noted above, have dashed earlier hopes.

A second area of regulatory change concerns data paths or the "transportation" of information. As competition in data transmission by wire became feasible, the Bell system responded by proposing Telpak, a "quantity discount" approach which was found to be discriminatory by the FCC and courts, see 31 FCC 2d 674, 3 CLSR 449. Specialized common carriers like MCI have attempted to "skim the cream" (to use the Bell expression) from the data transmission market by providing fast, economical, and efficient microwave transmission from one major point to another major point. Other transmission methods have been considered frequently by the FCC, e.g., cable television and domestic satellites. Lasers, waveguides, and one way multipoint distribution services will become important communications services in the near future. Bell, General Telephone and the independent telephone companies refuse to take these new competitors lying down. State and federal regulatory agencies must reevaluate existing regulations, both to recognize the new technology and to permit this technology to be used in the communications services at reasonable rates.

AT&T and others have begun a serious effort to encourage the state and federal regulatory commissions to continue the monopolistic approach. AT&T has recently filed a request that the FCC hold evidentiary hearings to consider the effects and future implications of "continued competition" rendered by newer specialized common carriers. In requesting the hearings, AT&T suggested that the FCC consider several major questions, including the degree to which the established common carriers have served the public interest over the years, the social and economic effects of the policy on existing common carriers and their customers, and the likelihood that bad experiences with regulated competition in other industries like the railroads and airlines will be repeated in the telephone industry. John

D. DeButts, Chairman of the Board of AT&T, recently urged the public utility regulatory commissions to declare a moratorium on "experimentation in economics" in order to consider the possibility that there may be sectors of our economy, telecommunications one of them, where the nation is better served by modes of cooperation than by modes of competition. Regulation regarding data transmission is changing and is bound to continue to change.

In 1973 packet communications has emerged as a technology which could be more efficient than current means of message switching especially for some types of computer communications. Message switching is currently the primary domain of communications companies because of the FCC's Computer Inquiry which dragged on from November 1966 to April 1973. Over 3000 pages of information were considered by the FCC and evaluated by the Stanford Research Institute, followed after further deliberation, by a tentative decision in 1970. After more responses from the industry were received, the FCC rendered a final decision in 1971 which, in turn, was appealed to the courts by the General Telephone Corporation and generally affirmed. Only after five and one-half years of testimony, rebuttal and judicial appeal, was a final decision reached. The length of time required for these deliberations is outrageous but not atypical. More of the same can be expected before the federal and state regulations regarding network regulation becomes clear.

CONCLUSIONS

What is the importance of the regulation of communications? Although 85-90 percent of communications now is voice traffic, by 1985 it is estimated that more than 50 percent of communications traffic will be data. Some of the major users of data transmission in the late 1980's will be colleges and universities. It is up to university administrators and computer users to have some familiarity with the regulatory problems involved. Whether one likes it or not, one is going to have to deal with the federal government, and with state governments, in order to get *economic* communications services for computing in higher education.

Bell System Services for Digital Data Transmission

by Carl F. Stuehrk
AT&T

It appears that Data Communications, both present and proposed, are affording exciting opportunities to the Data Processing Manager for improved utilization of resources with corresponding lower costs. Designers and administrators of sophisticated computation centers for colleges and universities are concerned both with being able to perform the offered assignments and with performing that work in the most efficient, economical manner. One must be keenly aware of the need to balance and optimize computer resources and data communications capabilities, and as such must be concerned with evaluating the various alternatives available. An increasing number of alternatives are available in the form of Private Line dedicated channels of varying speeds and capabilities, specialized as well as general networks, common user type services such as Dataphone WATS, and combinations thereof.

The Bell System has traditionally offered a variety of services both of a facility as well as a network nature. Today, in fact, the ubiquitous, nationwide message network still constitutes the largest, most flexible, and most heavily used data network in the world. While the Bell Systems does not currently offer a specialized PL type data switching network service, it does offer a wide variety of data communications services which enable the sophisticated user to construct operational data networks of one's own. AT&T expects to continue the role of a major provider of basic data communications services and will expand and augment services in the future to be responsive to the needs of the industry.

A new Bell System data communications service should afford higher education users even more alternatives in the future. Dataphone Digital Service represents a new environment for system planners, designers, and managers, and should provide new opportunities for designing and optimizing teleprocessing/computer networks.

For almost 15 years, data transmission has been provided generally on analog type facilities. These services require conversion to adapt the business machine signal format to the analog transmission medium. The modem converts the binary signal of the terminal into analog form for transmission over the communication facility, and at the end of the receiving end, performs the reciprocal function.

Analog systems have grown rapidly over the past decade, with increasing volumes of data being moved over the public switched network as well as over many private circuits and networks. Despite known design limitations and a recognition that digital transmission is inherently a more efficient medium for transporting digital data, analog data systems have performed well and have grown to a high degree of complexity and sophistication. The Bell System and others have devoted, and are continuing to devote, considerable effort to improved analog performance since analog channels are expected to occupy an important position in the data communication world for many years to come. The new Bell System family of modems recently introduced, including the 201C, 208A, and others yet to be announced, are concrete evidence of AT&T's continued development in this direction. A growing demand for even higher speeds, greater accuracy, and increased through-put however has dictated the need for a new digital data transmission environment.

In the new environment employing end-to-end digital transmission, signal conversion will not be necessary. Throughput capacity and efficiency should increase and total costs should go down with the elimination of the modem and the improved error performance. These goals are realized as a result of the inherent nature of the digital technology as well as the specific design of the system.

The potential for digital data systems was recognized by the Bell System as early as 1961 even as the first digital transmission systems were being installed to carry voice circuits in heavily congested metropolitan areas. While it was technically feasible to use the system for data, the economics of doing so did not support the idea. At that time, the data market was still in its infancy and data requirements were simply not of a magnitude sufficient to match the characteristics of what is basically a high capacity system. The volume could be handled very well by the existing analog plant, both from a technical and an economical standpoint. Since 1961, however, data communications has continued to grow at a rate of 25-30 percent per year. Accordingly, by 1970 it was felt that the timing was right, technically and economically, for the application of digital technology to data communications. Planning and design began for a new digital data system.

The application of digital technology by the Bell System takes shape as *Dataphone® Digital Service.* First and most important, it is a data only system, i.e., it does not provide voice coordination. Initially, service provided over DDS will be private line, full duplex in nature. It will offer synchronous transmission speeds of 2.4, 4.8, 9.6 and 56 kilobits per second. Subject to regulatory approvals, point-to-point private line service will be available in early 1974, and multipoint service will follow by mid-1974. The Bell System is exploring the possibilities of offering other versions of DDS as well as other transmission speeds above and below the initial speeds.

The DDS architecture consists of three different sub-systems. The *first* sub-system consists of the local facilities connecting the customer premises to the local telephone central office. The *second* includes the link between the local central office and the digital data hub office. The hub office and the nationwide intercity digital network make up the *third* sub-system.

At the customer premise, digital facilities will be terminated in the DDS interface. The signal on the line side of the interface is different than on the terminal side because the design includes a signal format that is bi-polar in nature between the DDS interface and the local central office. The DDS interface processes the transmitted signal into the bi-polar format and regenerates it in the process. The interface also provides an important testing capability with local and remote test centers. Because the digital signal is synchronous (timing is provided by the network), another important function of the interface is to provide timing recovery so that the business machine terminal is synchronous with the transmission medium. Two types of interfaces will be available: the data service unit (DSU) that provides all functions; and the channel service unit (CSU) that provides only channel terminating and testing capability. The availability of two types of interfaces provides an option to the terminal manufacturer to incorporate the three functions of timing recovery, signal processing, signal regeneration into its terminal design.

At the local central office a number of different customer services will be brought together, irrespective of channel speed. The central office also connects to the digital hub via a T-1 carrier line. At the central office, regeneration of each signal takes place and each separate channel is fitted into an assigned time slot of a T-1 carrier by multiplexing techniques. One T-1 system has the capability of handling 460 2.4kb channels, 230 4.8kb channels, 115 9.6kb channels or 23 56kb channels. Because multiplexing equipment in the local central office is flexible, it can mix and match the various speeds to provide optimum utilization of the T-1 system.

The DDS hub office terminates local facilities and connects these facilities to the intercity digital facilities. It is also the principal location for testing of all long haul and local distribution links. At the hub, a highly accurate timing source is maintained and synchronized with other timing sources throughout the network. In this way, precise network timing is generated which controls the transmission system within each local area. The hub office also performs functions similar to the local office, but higher up in the network hierarchy. It multiplexes local lines and T-1 lines into various types of long haul, high capacity facilities that make up the intercity digital network.

One of the newest of the high capacity facilities, "DUV" or Data Under Voice, is a transmission system in which a presently unused portion of the frequency spectrum on existing microwave radio channels is employed for the transmission of a 1.544mb data stream. Most Bell System microwave radio channels are designed to transmit up to 1800 voice channels divided into three master groups of 600 channels each. Since voice channels occupy a frequency spectrum that extends from about .6mhz to about .8mhz, the segment between 0 and .6mhz is generally unused. DUV takes advantage of this and impresses a 1.5mb data stream within this frequency slot. With this new development, there is the capability of providing digital data transmission anywhere radio facilities are available; which amounts to efficient utilization of existing resources.

DDS represents a new environment in which higher and greater reliability should be very possible. Concentration functions take place at the local central office, where many separate service channels of diverse speeds are combined into one common short haul T-1 carrier. A similar arrangement also takes place at the hub office. Channels from many T-1 lines and local lines are combined and multiplexed into common long haul digital facilities. At these two stages of multiplexing testing features have been engineered that will improve system performance and assure greater system reliability.

At the local central offices and at the digital hub offices, on-going monitoring of the quality of the service takes place. This performance monitoring looks for signal format violations and timing violations which, in the DDS system are prime indications of difficulty. If a difficulty is detected, the system will automatically switch to standby systems. Thus rapid restoration is assured if transmission problems occur. Similar monitoring and protection arrangements take place between digital hub offices. If an individual DUV channel becomes sub-marginal, it will automatically switch to a standby protection channel. In the DDS design, an envelope of protection envelopes the service from the transmitting local central office to the receiving central office at the other end of the circuit. Within the envelope of protection, monitoring and automatic switching will assure a high level of performance and reliability.

There are attractive features in the loop plant as well with testing capabilities built into each Hub office. A test person in Chicago or New York, for example, wherever the trouble is first noticed, has the ability to test the end of the circuit in one's own area at all critical points: at the local central office; at the line side of the station interface; and at the terminal side of the station interface. In addition, one will be able to initiate a loop around test at each of these points. In this way, the trouble can rapidly be isolated and the proper remedy determined. However, in addition to being able to detect trouble in the local circuitry, it is also possible to make identical tests at every station on the circuit.

All of this means better service in terms of lower error rates, minimal downtime and rapid restoration. Design objectives for DDS are to provide end-to-end transmission performance which will average no more than one error second in 200 seconds of transmission at 56 kilobits per second. It is expected that average circuit downtime should not exceed three to four hours per year!

While DDS represents a new environment, it is a system that lends itself quite readily to transition. It will be compatible with present analog services operating at the same speeds. The station interface is an EIA standard at 2.4, 4.8 and 9.6 kb and a CCITT interface at 56 kb. The data transmission format is such that the network is transparent to all data languages.

What about price? Digital technology represents a basically lower cost method of transporting digital data from one place to another. AT&T hopes to pass on to the users of these services the benefit of this inherent cost savings. The exact level of rates for service utilizing DDS, which are now under study preparatory to filings with the FCC and other regulatory agencies, will be shaped by several influences including the market and cross elasticities between other services both Bell and non-Bell. In October, 1972 AT&T filed 214 applications with the FCC to construct the initial intercity digital facilities for the 1974 network. This filing included "illustrative rates" which generally were lower than comparable analog rates. However, these rates are now being given extensive study and, subject to regulatory approvals, AT&T will file appropriate rates before the year is out.

When will Dataphone Digital Service be available? Subject to regulatory approval, service is expected to be available between New York and Boston in January 1974 and by May 1974 will be expanded to include Washington, Philadelphia and Chicago. According to the planned schedule for rapid and steady network growth from the east coast across the country, by the end of 1974, 24 metropolitan areas will be served on a coast to coast basis, and by the end of 1975, DDS will serve a total of 60 cities. By the end of 1976, 96 cities will be covered. Customers having requirements in cities not initially served by DDS, will be able to utilize analog facilities to connect to the Digital Data System. Plans are also underway to offer Intrastate and local DDS by the Bell Operating Companies on a coordinated basis with the interstate development. In summary, the Bell System believes that what data users really want is a communications capability that is relatively error-free, efficient, and economical, and easily integrated into their systems. DDS provides for this.

PCI's VANLINE Service

by Lee B. Talbert
Packet Communications, Inc.

To explain what the Proposed service of Packet Communications Inc. may mean to educational users, one should begin with an explanation of PCI's service philosophy. PCI believes that a mature data communications network service should be transparent to the user. A user should be able to use a communications outlet in much the same way that he uses electrical current. The average user of data communications does not need to know, nor does he necessarily care to know, how the data communications network operates. What the user is concerned with is transporting his data to and from the computer on demand, with the assurance that his data be delivered reliably, accurately, and inexpensively.

This service philosophy is not present in today's communications marketplace. If a user is to make the best use of existing data communications facilities and the new data channel facilities and services now available, one must become deeply involved in communications systems engineering and management. Users who venture into data communications projects are now generally responsible for:

- Acquiring the necessary transmission channels (lines).
- Working out interconnect arrangements.
- Monitoring the lines for errors and loss of service.
- Hiring hardware and software specialists.
- Purchasing or leasing equipment in addition to that provided by the carriers.

- Planning for backup facilities in the event of line or equipment failure.

This tends to leave one little time or money to improve the content of what is communicated. In short, the necessary solution on order of magnitude larger than the basic problem. What is needed is a transparent communications service which leaves the user free to concentrate on the reason for communicating, not on the act of communicating. Freeing resources in this way, has important implications for individuals and institutions interested in achieving: 1) a discipline for communicating; 2) a way to eliminate unnecessary duplication of information resources; and 3) a set of protocols and a program of user education which makes the achievement of steps 1+2 a reality.

VANS

The VAN Service, or "Value Added Network" Service, is a relative newcomer to the communications marketplace. To the user a Value Added Network Service should provide a flexible and transparent computer communications network without the normal management headaches. VAN Service is produced by adding equipment and ancillary services to the facilities and services which are available from existing common carriers. The resulting service fulfills the specialized communications needs of computer communications users. In January of 1973, Packet Communications Inc. requested authorization from the Federal Communications Commission to offer a national packet-switched computer communications service on a shared network basis. Since then three other companies have made such filings. No VAN Service now exists, but it is a certainty that such will not be the case for very long.

PCI's proposed Value Added Network Service will be different from what is available from the common carriers in three ways. First, additional services will be tailored to meet the specialized needs of data and computer communications users. Second, packet-switching gives users a virtual network which uses costly resources only on demand for data to be sent. Finally, the potential for resources and facilities sharing opens new economies and extends geographic coverage for users.

Packet's VANLINE NETWORK SERVICE customers will obtain a service which is different from a transmission channel between two fixed points similar to the way that steel is different from steel valves. Some of the significant advantages will be:

- *Single vendor responsibility* for the complex computer/communications network.
- *Access to a national network* which gives the user the possibility of easy addition or deletion of terminal access points.
- *Distance-independent pricing.*

- *Error detection and correction* (less than one undetected error per year).
- *Automatic alternate routing* in the event of a line or node failure. This routing allows automatic line load leveling which will tend to overcome network congestion in peak periods of transmission and increase efficiency. Thus lower costs to users.
- *Standard interfaces and protocols.*

Specific protocols for computers to terminals and computers to computers will allow communication between dissimilar equipment. Available services will include: 1) terminal speed and code conversion to network standards; 2) local echoing which allows full duplex operation over great distances without "sticky keyboard" delay; and 3) local terminal interfaces which are dial-up and/or private line, synchronous or asynchronous. Thus, Packet's service is designed to be a complete communications system. However, this is not the full extent of the advantages. A good solution to the problem of data communications is not that simple.

PACKET-SWITCHING

Packet's VANLINE Service is based on a new technology, packet-switching, which permits PCI to assume full responsibility for a client's data communications while at the same time increasing performance and lowering costs. This technology is message oriented but differs in important ways from today's more conventional message-switched systems.

Packet-switching is the technique for transmitting data which was incorporated by the Department of Defense Advanced Research Projects Agency (ARPA) in its research network. Unlike today's conventional circuit-switched channel communications system in which an entire circuit must be dedicated to a given burstor data stream for the duration of a conversation, a packet-switched data communication network transmits information by interspersing small messages on shared wideband lines rather than by dedicating lines to the exclusive use of a given data stream. In packet-switching, data streams are divided into small segments called packets which can contain up to approximately 1000 bits or 130 characters of data.

In PCI's proposed network packets are forwarded by a network of minicomputers called Packet Switching Processors (PSPs) over optimal routes across the network in much the same way that freight cars are routed from coast to coast over a network of interconnected railroad lines. By sharing the network among a great many users, the result is efficient use of dedicated wideband lines, and redundancy of available routes through the PSP minicomputers and their connecting lines.

Packet Communications Inc. will provide its proposed service with a physical plant composed of channel facilities leased from the common carriers and interconnected by mini-switching computer equipment. The actual network would consist of similar sections in many cities interconnected in a distributed fashion by inter-city trunk facilities. Several specific components will be provided by PCI: PSPs; TAPs; Network supervision; and Customer computing.

PSPs permit full two-way conversations among computers on the network. PSPs also support the distributed receive-and-forward function which is so fundamental to the packet-switching concept. Full error checking and retransmission is used on all network lines except those connecting to terminals which cannot support error control. The probability of an undetected error will be less than 10^{-12} through the use of a 24 bit cyclic redundancy checksum.

Certain of the computers connected to PSPs on the network will be operated by Packet Communications Inc. TAPs, for Terminal Access Processors, are one kind. TAPs provide for access to the network of a variety of terminal types of different speeds, codes, and models. The various data streams from these terminals are translated into a standard network internal format. TAPs then concentrate and multiplex these data streams through their serving PSP into the network and thence to the various computers specified by the terminal users. Terminal-to-TAP access can be by private line were justified, or more generally, by dial access line. Because these interconnections are almost identical to those provided by remote-access computing service companies today, terminal users will be able to use virtually any type of terminal and line in much the same way they do at present, with greater flexibility in terms of computers and services available and with greater reliability and ease of use, yet lower overall costs. It is important to observe that the TAPs provide no data processing in any common sense. Rather they provide format control, code translation, multiplexing, accounting for usage and other fundamental communications functions.

Another category of computers connected to the network is the Network Operations Center (NOC) computer. The functions of these centers are: accounting and billing; network status monitoring; and control of operations. There will be two such centers in the network, one in the Boston area, and the other in the western United States.

The final and most important category of network computers is the customer computer. Customer computers can be connected in a variety of ways to a serving PSP and thus to the network. These vary as to reliability, bandwidth, and degree of Packet responsibility for the connection.

What does transparency of network service mean? PCI proposes to assume the burden of acquiring lines, monitoring them, and maintaining them. In addition Packet Communications Inc. also will be responsible for the receive-and-forward strategy used to transmit a message without errors. The user need never be concerned with how TAPs communicate with PSPs or how PSPs communicate with PSPs. He or she does not need to know how the network works internally any more than one needs to know where or how the electricity used to start a terminal is being generated and transmitted to the outlet. What one does need to know is how to interface with the outlet so as not to be electrocuted. The same is true of the PCI network where the TAP is an outlet and certain procedures must be followed if communication is to be permitted.

The host or user computer can be connected to the PSP by the user or by Packet Communications Inc. If PCI's hardware and software are used, the user only needs to learn to use the network communication program which interfaces his computer to the network. It is important to note that the network does not involve users in the job of running the network. Rather, the user is freed to become involved in the decision processes which affect one's computer operation: operating system commands; resource sharing between computer centers; and special programs which allow information to be translated into another machine's code once it has been successfully communicated. All of these are concerned with what the user is doing. That is computing. Users can get down to the job of computing, or resource sharing once they are relieved of the burdens of communications. When they perceive that what Packet Communications Inc. is proposing is a virtual network scheme which gives reliable, cost effective communications on a demand basis, then new applications and new industries will be created to make use of this new national marketplace for computing and databased services.

CONCLUSION

Packet-switching combines advances in network design, distributed network control, and low-cost minicomputers, to provide a practical solution to a nationwide need for a national, public, high-performance data communications system. Packet-switching is well suited to new communications technology and to the opening of the marketplace. Packet-switching networks are capable of achieveing substantially more efficient line utilization, greater reliability, higher transmission speeds, and greater flexibility than heretofore has been realized by any available computer communications systems.

With a community of users sharing such a communication system, the user-interfacing minicomputers distributed through the network also may be used to provide user-interfaced protocol and, in effect, provide a common language for communication between differing types of equipment and data organizations. The result is a common marketplace for computing, for remote terminal access and for computer resource sharing.

Common protocols can be provided for access from a variety of terminals (interface, batch, intelligent, transaction, etc.) to any of a variety of computer makes and types, for computer load sharing, and so on. A wide range of user data speeds, as well as such problem traffic as bursty transmissions and short transaction messages may be intermixed without loss of network efficiency. A further byproduct of the multiply-connected packet-switched network is that, once a given geographic area is spanned, total network cost varies little with traffic distance and thus true geographically-independent pricing is practical. Such characteristics make packet-switching uniquely well-suited to a publicly-shared data communications service.

What does this all mean for members of EDUCOM? Only the future will supply definite answers. However, certain facts are apparent:

- A value added network will enable users to get down to the business of computing.
- Such networks will make resource-sharing on a national basis economically viable.
- There are economies of scale which could bring down the cost of library services, paper, special education programs.
- The existence of packet-switched computer communications service will make possible small research grants devoted to user protocols and manuals which teach individuals how to share resources.
- A public network might overcome some of the political problems within a state where the legislature feels it might be good for the police and universities and highway departments to share the same facilities. Each facility would pay only for what it used and would have clear accounting for that use.

In the end each customer must evaluate the packet approach in the light of the amount and types of information one sends and the number and kinds of political problems a public network will solve. A public network can save substantial amounts in equipment and software rented. It can eliminate the need to change equipment. It can ease accounting and billing loads, and it can facilitate computer communication between entities who prefer to remain autonomous.

Appendix A

FIND System Commands

Elementary Commands

ATTRIBUTES	List attributes in any data base
EXPLAIN	Explain a command
RETRIEVE	Sets up working data base for a session with FIND
SELECT	Extracts specified subset of working data base
SORT	Re-orders working data base by specified criteria
PRINT	Displays the working data base on the terminal
STATISTICS	Computes statistical measures
CANCEL	Halts a command
STOP	Terminates a session with FIND

Advance Commands

CORRELATION	Print correlation between two or more attributes
DEFINE	Composes new attributes based on old ones
EXECUTE	Performs FIND commands stored in a file
FIT	Provides linear and exponential fit routines
FORMAT	Allows flexible formatting of output
OUTPUT	Allows output to a saved file
RETRIEVE	Additional information on use of the RETRIEVE command
SORT	Additional information on use of the SORT command
XTAB	Provides frequencies and cross-tabulations

Miscellaneous Commands

COUNT	Count attributes and entities
LABEL	Insert textual information into FIND output
LOAD	Retrieve a previously saved working data base
REDUCE	Permanently shrink a working data base
RENAME	Change the name of an attribute in the working data base
RESTORE	Restore the working data base to original state
SAVE	Save a copy of the working data base
SCRATCH	Destroy a portion of the working data base
TIME	Print running time

Data Base Maintenance Commands

UPDATE	Initiates a data base update session
IDENT	Specifies a list of identifying attributes
LOG	Specifies a file in which to record changes
EXIT	Terminates a session with UPDATE
ALTER	Specifies data items for immediate modification
MODIFY	Specifies a list of attributes for deferred modification
ADD	Creates a new entity
DROP	Deletes an entity from the data base

Appendix B

FIND Available Data Bases
(December 1973)

- Administrative
- Faculty
- Staff

Personnel

- Departmental Budgets & Expenditures
- Student Courses
- Physical Facilities
- Art Gallery Collection
- Alumni (Development Office)
- Endowment Funds

Appendix C

Resident File OMIC, 1973

PRIMARY

PAGE 1

FILE NAME	NUMBER OF RECORDS	– SEGMENTS –			NUMBER OF ELEMENTS	AVG. REC. LENGTH	TOTAL NO. BYTES	SPACES USED
		TYPES POSS.	TYPES USED	TOTAL USED				
ACADEMIC PROGRAM SUPPORT	588	8	8	2,078	105	246	144,740	177
ASSIMILATED STUDENT PROFILE	27,583	65	38	551,757	414	651	17,946,358	17,940
A.S.P. – TEST FILE	921	65	37	18,149	109	707	651,010	833
COURSE SECTION MASTER	5,889	14	4	13,336	307	94	553,168	491
PERSONNEL/PAYROLL DIST.	4,484	45	30	70,895	307	410	1,838,218	2,070
P.P.D. – TEST FILE	444	45	35	7,620	307	590	262,040	322
FACILITIES ROOM INFORMATION	9,984	8	2	16,590	53	55	544,692	1,046

NUMBER OF ELEMENTS DOES NOT INCLUDE INDIRECT REFERENCES OR GROUP NAMES.

TOTAL NUMBER OF CHARACTERS DOES NOT INCLUDE RECORD GROWTH FACTORS OR SLACK BYTES.

SPACE USED IS CALCULATED IN 'PAGES'. A PAGE = 1,692 BYTES, FOUR (4) PAGES FIT ON A TRACK OF A 2314. THE SPACE USED DOES NOT INDICATE SPACE ALLOCATED TO PERMIT FOR ADDITIONAL INFORMATION' (E.G., RECORD EXPANSION AND/OR ADDITION.)

SECONDARY

FILE LABEL	NUMBER OF RECORDS	– SEGMENTS –			NUMBER OF ELEMENT	AVG. REC. LENGTH	TOTAL NO. BYTES	SPACES USED
		TYPES POSSIBLE	TYPES USED	TOTAL USED				
ACADEMIC STATUS	4	1	1	4	2	16	64	4
ACHIEVEMENT TEST NAMES	16	1	1	16	2	18	288	4
AGENCY NAMES	363	1	1	363	2	44	15,972	19
AID INTENT DESCRIPTION	7	1	1	7	2	28	196	4
COMPUTER BILLING CHARGES	56	1	1	56	3	40	2,240	5
BUILDING FILE	316	5	3	911	23	113	35,612	44
CAMPUS NAMES	8	1	1	8	2	18	144	4
CLASS NAMES	17	1	1	17	3	26	442	4
COLLEGE NAMES	39	1	1	39	3	26	1,014	4

FILE LABEL	NUMBER OF RECORDS	– SEGMENTS –			NUMBER OF ELEMENTS	AVG. REC. LENGTH	TOTAL NO. BYTES	SPACES USED
		TYPES POSSIBLE	TYPES USED	TOTAL USED				
BUILDING COMPLEX TABLE	3	1	1	3	2	16	48	4
DEGREE NAMES	19	1	1	19	2	28	532	4
DEPARTMENT NAMES	210	1	1	210	3	30	6,300	13
DEPARTMENT CODES	171	1	1	171	4	32	5,472	12
DISCIPLINES	0	1	0	0	3	30	0	0
JOB CODE WORK FILE	549	1	1	549	2	32	17,568	25
ROOM FUNCTION TABLE	21	1	1	21	2	16	336	4
HIGH SCHOOL NAMES	2,361	1	1	2,361	2	26	61,386	85
JOB CODES	550	1	1	550	2	32	17,600	24

SECONDARY

FILE LABEL	NUMBER OF RECORDS	– SEGMENTS –			NUMBER OF ELEMENTS	AVG. REC. LENGTH	TOTAL NO. BYTES	SPACES USED
		TYPES POSSIBLE	TYPES USED	TOTAL USED				
LANGUAGE NAMES	25	1	1	25	2	14	350	114
LEASE TABLE	4	1	1	4	2	12	48	4
MAJOR NAMES	201	1	1	201	6	40	8,040	11
RACE NAMES	5	1	1	5	2	12	60	4
REPORT WRITER FILE	45	11	10	569	1	3,184	143,292	111
STUDENT STATUS DESCRIPTION	17	1	1	17	2	24	408	4
RACE SUBGROUP NAMES	4	1	1	4	2	18	72	4
VERMONT TOWN NAMES	248	1	1	248	2	26	6,448	10
FACILITIES TYPE TABLE	92	1	1	92	2	13	1,656	5
FACILITIES USABILITY TABLE	5	1	1	5	2	16	80	4
TOTALS (4 PAGES)					**1,796**		**22,265,894**	**23,303**

Appendix D

Sample Queries of University of Vermont Data Bases

EXAMPLE 1.

Find the overall percent of salary increase for the University for FY '74.

KEY: Pay.Type - 1 and 2, Salaried Employees
LSTA.Code - Leave,Sabb,Terminated,Active
Current.Sal.Base - Current Fiscal Salary
Pri.Sal.Base - Prior Fiscal Salary

STEP 1: Sum Cur.Sal.Base — Sum Pri.Sal.Base
Where Pay.Type RN 1/2 and LSTA.code NE 'T' and Pri.Sal.Base NE 0.

RETURNS: Expression
Sum Cur.Sal.Base Value 21,885,265.36
Sum Pri.Sal.Base Value 20,658,168.00

1,227,097.36

STEP 2: 100 * 1227097 / Sum Pri.Sal.Base
Where Pay.Type RN 1/2 and LSTA.Code NE 'T' and Pri.Sal.Base NE 0.

RETURNS: Expression
Sum Pri.Sal.Base Value 20,658,168.00

5.94

EXAMPLE 2:

STEP 1: Records Count H.rank max H.rank avg H.rank min H.rank where SY Eq 373 and Fin Eq 'A' and Sex Eq 'F' and Res Eq 'I'.

RETURNS:	Record count is	880	# in select
	Count H.rank value	760	# had H.rank inputed
	Max H.rank value	80	
	Avg H.rank value	60	
	Min H.rank value	25	

Population is in-state, females who have applied and been accepted for Fall '73 semester.

H.rank is used as a standard score scale for prediction studies of each year's applicant population.

EXAMPLE 3:

STEP 1: Sum Amt.Awarded.A Where College.Dept Eq 'Medicine' and Award.Start.Date.A Gr 720630 and Award.Stop.Date.A Ls 730701.

RETURNS: Gives $ amount for Dept of Medicine for all awards that started and ended within the fiscal year.

STEP 2: Project.Title Amt.Awarded.T Start.Date.T Stop.Date.T Where Director Eq Smith, TH

RETURNS: Gives all projects, total amount awarded for each and date of period covered for one faculty member.

STEP 3: Applic.Type.A Award.Stop.Date.A Spon.Code UVMID Director where Award.Stop.Date.A RN 730901/731231

RETURNS: Gives all current awards that will run out within the Sept-Dec. period of this year.

STEP 4: Sum Tot.Dols.Req Where Applic.Date.N RN 730101/730917 Tand Pend. Code.N Eq 'P' and Spon.Code RN 'AN0000'/'AN9999'.

RETURNS: Gives total $ amount requested from National Science Foundation since the 1st of the year, which has not yet been awarded or rejected (is still pending) — new applications, not additions to old ones.

Appendix E

Agenda SIGSOC Working Session on COMPATIBILITY

1:30 – 5:30, Wednesday, June 6, 1973, in the
Regency Flair Meeting Room, Americana Hotel
7th Avenue at 52nd Street, New York City

SESSION CHAIRMAN:	G. R. Boynton, Political Science Program Director, National Science Foundation
DISCUSSANTS:	Hugh Cline, Russell Sage Foundation George Sadowsky, The Urban Institute

The SIGSOC Working Session on Compatibility attempts to summarize the current situation and then develop some consensus on steps to be taken to improve the situation in the future. The last hour of the session will be devoted to discussing and voting on a series of specific resolutions.

PROGRAM

1:30 "An overview of social science computing and some recent efforts relevant to compatibility"
R. Anderson, SIGSOC Chairman

1:45 "CONDUIT project findings pertaining to compatibility"
T. Dunnagan, CONDUIT Central, Duke University

2:00 "Design of an optimally compatible social data analysis system"
E. Coover, J. Dyer, J. Gross, C. Johnson, J. Lutgen, J. Miner, University of Minnesota

2:20 "A preprocessor permitting P-STAT to be written for five different computers"
R. Buhler, Princeton University

2:40 "The approach to compatibility of the Cambridge Project and the Consistent System"
J. Klensin, D. Yntema, R. Wiesen, J. Markowtiz, Massachusetts Institute of Technology

3:00 "Data and software compatibility problems: the experience of the ICPR"
G. Marks, ICPR, University of Michigan

3:20 "Data structures and codebook formats: the Berkeley experience"
M. Shanks and H. Weinstein, Survey Research Center, University of California at Berkeley

3:40 "SPSS plans relevant to compatibility"
K. Steinbrenner, NORC, University of Chicago

3:45 Coffee break

[illegible]:00 Discussants and Discussion

Appendix F

Conference Participants

Leon Adkison
Assistant Professor of Information Systems
Virginia Commonwealth University
101 South Floyd Avenue
Richmond, Virginia 23220

R. Albano
Associate Professor
New York City Community College
300 Jay Street
Brooklyn, New York 11201

Donald L. Alderman
Associate Research Psychologist
Educational Testing Service
Princeton, New Jersey 08540

Howard W. Alexander
Professor of Mathematics
Earlham College
Richard, Indiana 47374

Williard F. Allen
Associate Vice President
University of Alberta
Room 3-12, University Hall
Edmonton, Alberta, Canada

John E. Alman
Director, Computing Center
Boston University
111 Cummington Street
Boston, Massachusetts 02215

Mohammed M. Aman
Chairman, Department of Library Science
St. John's University
Jamaica, New York 11439

Ernest J. Anastasio
Associate Director (ODAR)
Educational Testing Service
Princeton, New Jersey 08540

Gary D. Anderson
Director, Health Sciences Computation Unit
McMaster University
Hamilton, Ontario, Canada

Ronald E. Anderson
Associate Professor
University of Minnesota
2122 Riverside Avenue
Minneapolis, Minnesota 55404

R. P. Andres
Associate Professor
Princeton University
Department of Chemical Engineering
Princeton, New Jersey 08540

Thomas Archibald
Chief, Management Information Systems
ODCSCOMPT
U. S. Military Academy
West Point, New York 10966

Rosemarie Arcieri
EDUCOM
Princeton, New Jersey 08540

Donald R. Arnold
Director, Office of Management Systems
New Jersey Department of
Higher Education
225 West State Street
Trenton, New Jersey 08625

Robert L. Ashenhurst
Director, Institute for
Computer Research
The University of Chicago
5640 South Ellis Avenue
Chicago, Illinois 60637

Daniel M. Asquino
Staff Associate (Systems Development)
Massachusetts Board of Regional
Community Colleges
177 Milk Street
Boston, Massachusetts 02109

Dr. William F. Atchison
Acting Chairman, Computer Science
Department
University of Maryland
College Park, Maryland 20742

Hugh C. Atkinson
Director of Libraries
The Ohio State University
1858 Neil Avenue
Columbus, Ohio 43212

R. G. Auchincloss
Education Industry Research Manager
IBM Corporation
Department B41, 702 Building
Post Office Box 390
Poughkeepsie, New York 12603

Charles P. Austin
Assistant Director
West Virginia University Computer Center
Morgantown, West Virginia 26806

John E. Austin
Associate Director
Office for Information Technology
Harvard University
Cambridge, Massachusetts 02138

Richard B. Bagby
Director, Office for Computer Services
Rochester Institute of Technology
One Lomb Memorial Drive
Rochester, New York 14623

J. P. Balabon
Operations Manager
Mercer County Community College
Box B 1200 Old Trenton Road
Trenton, New Jersey 08690

David G. Bamford
Executive Assistant to the Vice President
University of Saskatchewan
Saskatoon, Saskatchewan, Canada S7N OWO

William P. Barley
Education Industry Marketing
IBM Corporation
10401 Fernwood Road
Bethesda, Maryland 20034

David G. Barlow
Assistant to the Controller
Harvard University
Room 458-1350 Massachusetts Avenue
Cambridge, Massachusetts 02138

Dennis S. Bartow
Associate Dean of Faculty,
Science/Technology
Mercer County Community College
1200 Old Trenton Road, Box B
Trenton, New Jersey 08690

Charlie Bass
Assistant Professor
Computer Science Department
University of California
Berkeley, California 94720

Robert M. Bassford
Planning Coordinator
Department of Budget and Fiscal Planning
State of Maryland
Treasury Building
Annapolis, Maryland 21404

George W. Baughman
Director, Special Projects
The Ohio State University
2130 Neil Avenue
Columbus, Ohio 43210

James C. Baxter
Senior Director, Computing Affairs
The University of Wisconsin System
Post Office Box 912
Madison, Wisconsin 53701

Barry Beckerman
Manager, Information Services
UCLA Campus Computing Network
405 Hilgard Avenue
Los Angeles, California 90024

Stoughton Bell
Director, Computing
University of New Mexico
Albuquerque, New Mexico 87131

D. H. Bent
Director, Computing Services
University of Alberta
Edmonton, Alberta, Canada

Daniel L. Bernitt
Pennsylvania State University
University Park, Pennsylvania 16802

John A. Bielec
Assistant to Vice Chancellor
University of Maryland
College Park, Maryland 20742

Julian Bigelow
Institute for Advanced Study
3 Horner Lane
Princeton, New Jersey 08540

Robert Bigelow, Esq.
28 State Street, Room 2200
Boston, Massachusetts 02109

Milton Birnkrant
Stevens Institute
Castle Point Station
Hoboken, New Jersey 07030

Robert P. Blanc
Project Manager
National Bureau of Standards
Tech B216
Washington, D.C. 20234

Thomas Blaskovics
West Virginia University
509 Forestry Tower Addition
Morgantown, West Virginia 26506

Taveta K. Bogert
Systems Programmer
National Board of Medical Examiners
3930 Chestnut Street
Philadelphia, Pennsylvania 19073

Ronald W. Boling
Director of University Computer
Services Planning
The University of Tennessee
206 Administration Building
Knoxville, Tennessee 37916

Jean C. Bonney
Acting Assistant Director, CCIS,
Business and Information Services
Rutgers Hill Center, Busch Campus
New Brunswick, New Jersey 08903

Kenneth Boutwell
Vice Chancellor for Administrative Affairs
State University System of Florida
Tallahassee, Florida 32306

Kenneth L. Bowles
Computer Center
University of California, San Diego
Post Office Box 109
LaJolla, California 92037

Ronald W. Brady
Vice President for Planning
and Allocation
University of Illinois
349 Administration Building
Urbana, Illinois 61801

John Brandon
Assistant Professor Computer Science
U. S. Merchant Marine Academy
Kings Point, New York 11024

Thomas Braun
Coordinator of Information Systems
Council on Public Higher Education –
Kentucky
Capital Plaza Office Tower
Frankfort, Kentucky 40601

Hunter M. Breland
Associate Research Psychologist
Educational Testing Service
Princeton, New Jersey 08540

Jane Breyer
Abacus Associates
443 Park Avenue, South
New York, New York 10022

Charles A. Brooks, Jr.
Director – MIS Computerization
South Carolina Commission on
Higher Education
1429 Senate Street
Columbia, South Carolina 29201

Leonard Brush
Assistant Director for Information
Systems
212 Engineering C
The Pennsylvania State University
University Park, Pennsylvania 16802

Lee T. Bryant
Assistant Professor
State University of New Yo[illegible] Geneseo
Geneseo, New York 14454

James R. Buchholz
Vice President Business and Finance
Rochester Institute of Technology
One Lomb Memorial Drive
Rochester, New York 14423

Roald Buhler
Computer Center
Princeton University
Princeton, New Jersey 08540

Robert E. Burns
Associate Dean
College of Arts and Letters
University of Notre Dame
Notre Dame, Indiana 46556

Russell W. Burris
Director, Consulting Group on
Instructional Design
University of Minnesota
Minneapolis, Minnesota 55455

W. Richard Burrows
Associate Director, Administrative
Systems and Computational Services
University of New Hampshire
Thompson Hall
Durham, New Hampshire 03824

Thomas E. Byther
Academic Coordinator
Computing and Data Processing Services
University of Maine
Wingate Hall
Orono, Maine 04473

Gerry Campisi
Computer Center
Educational Testing Service
Princeton, New Jersey 08540

Kevin Carolan
Associate Dean
Marist College
Poughkeepsie, New York 12601

Donald W. Carroll
Dean, The Wharton School
University of Pennsylvania
Philadelphia, Pennsylvania 19174

George A. Carroll
Director of Management Affairs, Library
Rutgers University
New Brunswick, New Jersey 08901

Richard I. Carter
Director, Computation Center
Northeastern University
Boston, Massachusetts 02115

Pierre Charlebois
Assistant to Director, Computing Center
Universite du Quebec
P. O. 500
Trois-Rivieres, Quebec, Canada

Henry Chauncey
President
EDUCOM
Princeton, New Jersey 08540

Richard Cheng
Acting Department Chairman
Computer Science
Rochester Institute of Technology
Rochester, New York 14623

Robert L. Chew
Director of Graduate Student Services
University of Massachusetts
Amherst, Massachusetts 01002

James Chlipala
Graduate Student – Design Engineer
Academic Computing Center
University of Vermont
Burlington, Vermont 05401

David L. Clark
Director, Computer Center
University of Arizona
Tucson, Arizona 85721

John A. Clark
Director, Office of Data Processing
Management
New Jersey Department of Higher
Education
10 Bruck Park Drive
Trenton, New Jersey 08618

Robert Clark
University of Tennessee
Director Data Processing
Knoxville, Tennessee 37916

John Clymer
Bell Laboratories
Holmdel, New Jersey 07733

Bill Coldrick
Honeywell Information Systems
1615 Northern Boulevard
Manhasset, New York 11030

James B. Conklin, Jr.
NERDC, 233 SSRB
University of Florida
Gainesville, Florida 32611

Albert E.Conord
Director, Academic Computer Center
U. S. Naval Academy
Annapolis, Maryland 21402

John Cornelius
University of California, San Diego
Chemistry Department
3116 Vrey Hall
La Jolla, California 92037

Donald F. Costello
Director, Lincoln Computing Facility
University of Nebraska
Room 225 Nebraska Hall
Lincoln, Nebraska 68508

Frances Craig
Information Systems, Coordinator
Carnegie-Mellon University
5000 Forbes
Pittsburgh, Pennsylvania 15213

Ronald D. Crain
Visiting Lecturer
Chemistry Department
University of Kansas
Lawrence, Kansas 66045

Arthur O. Cromer
Instruction and Research Manager
University of Louisville
Computer Center
Louisville, Kentucky 40208

Michael P. Cruskie
Information System Specialist
New York State Education Department
Room 1923, Twin Towers
99 Washington Avenue
Albany, New York 12210

Effa Dalton
EDUCOM
Princeton, New Jersey 08540

D. Derwood Davis
Director, Data Processing
University of Pennsylvania
Frankling Building
3451 Walnut Street I6
Philadelphia, Pennsylvania 19174

Thomas E. Davis
Assistant Professor of Operative Dentistry
New Jersey Dental School
201 Cornelison Avenue
Jersey City, New Jersey 07304

William I. Davisson
Professor of Economics
University of Notre Dame
Notre Dame, Indiana 46556

Ardath P. Day
Manager, Computer Center
Rutgers University
92 New Street
Newark, New Jersey 07102

Richard DeGennaro
Director of Libraries
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Con Dietz
Director, Computer Center
University of North Dakota
Grand Forks, North Dakota 58201

Norma Jean Dininger
Programmer Analyst
College of Medicine and Dentistry
Rutgers Medical School
Box 101
Piscataway, New Jersey 08854

Mary Mark Doubet
Director of Educational Research
and Information Systems
Villa Maria College
2551 West Lake Road
Erie, Pennsylvania 16505

Jane Draper
Associate Director
Center for Research in Computing
Technology
Harvard University
33 Oxford Street
Cambridge, Massachusetts 02138

Noreen Dulz
Trinity College
Hartford, Connecticut 06106

Karen A. Duncan
Assistant Professor
Medical University of South Carolina
80 Barre Street
Charleston, South Carolina 29401

Donald A. Dunn
Professor and Associate Chairman
Engineering-Economic Systems Department
Stanford University
Stanford, California 94305

Trinka Dunnagan
Senior Systems Analyst
CONDUIT
Duke University
Durham, North Carolina 27705

Harry A. Eick
Professor of Chemistry
Michigan State University
East Lansing, Michigan 48823

R. G. Ellis
Director, Computer Center
University of Pittsburgh
600 Epsilon Drive
Pittsburgh, Pennsylvania 15238

Henry Fallon
Princeton University
Forrestal Campus, Plasma Physics
Laboratory
Princeton, New Jersey 08540

David J. Farber
Information and Computer Science
Computer Sciences Building
University of California
Irvine, California 92664

Dennis Fife
Institute for Computer Sciences and
Technology
National Bureau of Standards
Washington, D.C. 20234

Rex Franciotti
Director, Computer Center
State University of New York
Stony Brook, New York 11790

Gene F. Franklin
Associate Provost for Computing
Stanford University
Durand 131
Stanford, California 94305

Howard Frazier
Princeton University
Forrestal Campus
Geophysical Fluid Dynamics Laboratory
Princeton, New Jersey 08540

Joseph B. Frechen
Associate Professor
St. John's University
Department of Math and
Computation Science
Jamaica, New York 11439

Greydon C. Freeman
Director, Yale University Computer
Center
175 Whitney Avenue
New Haven, Connecticut 06520

Eugene A. Fucci
Assistant Director
Kiewit Computation
Dartmouth College
Hanover, New Hampshire 03755

Elemer R. Gabrieli
Director, Clinical Information Center
School of Medicine, State University
of New York at Buffalo
E. J. Meyer Memorial Hospital
Buffalo, New York 14215

Alfonso Gandica
Information and Systems Sciences
Coordinator
Stockton State College
Pomona, New Jersey 08240

Mr. Gbago
West African Examination Council
Post Office Box 917
Accra, Ghana, West Africa

Martell J. Gee
Director, Office for Computing
Activities
University of Dayton
300 College Park Avenue
Dayton, Ohio 45469

R. Gene Geisler
Chairman, Timeshare Long Range
Planning Task Force
California State University and
Colleges
San Francisco, California 94132

Robert J. Gettelfinger
Executive Director, Data Processing
Educational Testing Service
Princeton, New Jersey 08540

Shirley Gilbert
Programmer Social Science User Service
Princeton University Computer Center
Princeton, New Jersey 08540

Walter J. Gilbert
Acting Director of Administrative
Computing Center
University of Maryland
Room 0127, Main Administration Building
College Park, Maryland 20742

Mark Gillen
Honeywell Information Systems
1615 Northern Boulevard
Manhasset, New York 11030

Philip Goldstein
Jersey City State College
Jersey City, New Jersey 07305

Raymond Gontasz
Honeywell Information Systems
1615 Northern Boulevard
Manhasset, New York 11030

Theresa Green
Coordinator of Program DELTA
University of Delaware
Newark, Delaware 19711

Timothy Griffin
Illinois Department of Finance
160 North LaSalle
Chicago, Illinois 60601

Alec Grimison
Associate Director, Office of
Information Systems
University of Puerto Rico
Post Office Box 4984-G
San Juan, Puerto Rico 00936

Stuart Grout
Director of Academic Planning
Boston University
116 Bay State Road
Boston, Massachusetts 02215

Richard E. Grove
Director, University Computer Center
Virginia Commonwealth University
1015 Floyd Avenue
Richmond, Virginia 23220

John Gwynn
Manager, Data Processing
University of Wyoming
Box 3945, University Station
Laramie, Wyoming 82071

Dale J. Hall
Director Computer Systems
Development
Computers and Information Services
Indiana University
Bryan Hall 22
Bloomington, Indiana 47401

Harold L. Hall
Director, University Computing Activities
University of Louisville
Louisville, Kentucky 40208

Keith A. Hall
Director of CAI Laboratory
The Pennsylvania State University
201 Chambers Building
University Park, Pennsylvania 16802

Harold C. Hannaway
Manager Marketing Support
Honeywell Information Systems
570 Taxter Road
Elmsford, New York 10523

Gustave A. Harrer
Director of Libraries
University of Florida
Gainesville, Florida 32611

Fred H. Harris
Director, Computation Center
University of Chicago
5640 South Ellis
Chicago, Illinois 60637

Lucille Heller
Institutional Research Associate
Rutgers, The State University
New Brunswick, New Jersey 08903

Carl Helm
City University of New York
Graduate Center
33 West 42nd Street
New York, New York 10036

Lyell D. Henry, Jr.
Assistant Vice President for
Educational Development and Research
University of Iowa
Old Capitol
Iowa City, Iowa 52242

Saul Herner
President
Herner & Company
2100 M Street, N.W.
Washington, D.C. 20037

E. C. Hertzler
Acting Associate Dean
The University of Michigan
4901 Evergreen Road
Dearborn, Michigan 48128

Freeman Holmer
Vice Chancellor for Administration
Oregon Department of Higher Education
Post Office Box 3175
Eugene, Oregon 97403

Robert P. Hopmann
Dean of Administration
Concordia Teachers College
River Forest, Illinois 60305

Priscilla Houston
Manager of Computer Services
Rice University
Institute of Computer Services
and Applications
Houston, Texas 77001

W. Carl Jackson
Dean of University Libraries
Indiana University C2 Library
Bloomington, Indiana 47405

Peter Jargowsky
Senior Programming Analyst
Educational Testing Service
Princeton, New Jersey 08540

William H. Jenkins
Director, Computing Centre
Queen's University
Kingston, Ontario, Canada

Michael Jennings
Director, Data Services Division
Oregon Department of Higher
Education
Post Office Box 3175
Eugene, Oregon 97403

George L. Jensen
Director
67 Administration Building
University of Tennessee
Knoxville, Tennessee 37919

James W. Johnson
Director, CONDUIT
University of Iowa
Iowa City, Iowa 52242

Jeremy E. Johnson
Director, Computing and Data
Processing Services
University of Maine
Wingate Hall
Orono, Maine 04473

Richard R. Johnson
Program Manager
Exxon Education Foundation
111 West 49th Street
New York, New York 10020

E. Walton Jones
Associate Vice President for Research
and Public Service Programs
University of North Carolina
Chapel Hill, North Carolina 27541

William J. Jones
Director, Academic Computing Center
Syracuse University
Syracuse, New York 13210

Edward D. Jordan
Director, Institutional Research
and Planning
Catholic University
Washington, D.C.

Andrew Kasarda
Center for Information Sciences
Lehigh University
Bethlehem, Pennsylvania 18015

I. Katsriku
West African Examination Council
Post Office Box 917
Accra, Ghana, Africa

Eugene J. Keenoy
Director, Computer Services
Kean College of New Jersey
Morris Avenue
Union, New Jersey 07083

William B. Kehl
Director, Campus Computing Network
University of California
405 Hilgard Avenue
Los Angeles, California 90024

Donald W. Keillor
Michigan State University
East Lansing, Michigan

Carl A. Kellogg
Data Processing Manager
State University of New York at Geneseo
Geneseo, New York 14454

Phyllis Kent
Assistant Director for Academic
Computer Services
Syracuse University Computing Center
Syracuse, New York 13210

H. Eugene Kessler
Executive Director
EDUCOM
Princeton, New Jersey 08540

Frederick G. Kilgour
Director
The Ohio College Library Center
1550 West Henderson Road
Columbus, Ohio 43220

Kenneth M. King
Dean for Computer Systems
City University of New York
New York, New York 10024

Ethel Kinney
EDUCOM
Princeton, New Jersey 08540

Dorothy L. Klein
Assistant to Director, Institutional Research
University of Connecticut
Gulley Hall, U-135
Storrs, Connecticut 06268

Dan Krevere
Educational Specialist
Sperry UNIVAC
1701 North Ft. Myer Drive
Arlington, Virginia 22209

Hyman W. Kritzer
Asst. Provost and Director of Libraries
Kent State University
Kent, Ohio 44242

Arthur Kroll
Guidance Programs
Educational Testing Service
Princeton, New Jersey 08540

E. Rex Krueger
Director, Computing Center
University of Colorado
Boulder, Colorado 80302

Thomas E. Kurtz
Director, Kiewit Computation Center
Dartmouth College
Hanover, New Hampshire 03755

Carolyn Landis
Secretary
EDUCOM
Princeton, New Jersey 08540

John F. Langdon
Vice President for Administration
Loyola University of Chicago
820 North Michigan Avenue
Chicago, Illinois 60611

Robert Langridge
Department of Biochemical Sciences
Princeton University
Princeton, New Jersey 08540

Richard LaRue
Director of Computer Center
University of South Dakota
Computer Center
Vermillion, South Dakota 57069

W. E. Lavery
Executive Vice President
Virginia Polytechnic Institute and
State University
Blacksburg, Virginia 24061

Samuel A. Lawrence
Vice President for Administration
Cornell University
303 Day Hall
Ithaca, New York 14850

Gerald R. Lawton
Assistant Director, Computer Systems
Development
State University of New York
99 Washington Avenue
Albany, New York 12210

Gerald Lazorick
Director, Mechanized Information Center
Ohio State University
Columbus, Ohio 43210

Mads Ledet
Director, Research and Education
Computer Service
University of Oklahoma
Norman, Oklahoma 73069

Michael E. Lesk
Bell Laboratories
600 Mountain Avenue
Murray Hill, New Jersey 07974

Priscilla Linsley
Library
Educational Testing Service
Princeton, New Jersey 08540

Jacques H. Loeliger
Assistant Director
Student Services Development
Rutgers University
New Brunswick, New Jersey 08903

Charles R. Loepkey
Director, Computer Center
University of California
Santa Barbara, California 93106

Robert A. Loscher
Director, Computer Services
Glassboro State College
Bole Administration Building
Glassboro, New Jersey 08028

Robert James Low
Vice President for Administration
Portland State University
Post Office Box 751
Portland, Oregon 97034

David Loye
EDUCOM
Princeton, New Jersey 08540

John F. Lubin
Associate Dean
The Wharton School
University of Pennsylvania
Dietrich Hall
Philadelphia, Pennsylvania 19174

Arthur Leuhrmann
Director, Project COMPUTe
Dartmouth College
Kiewit Computation Center
Hanover, New Hampshire 03755

David S. Macey
Coordinator-Office of Computer
Coordination
Council of Ontario Universities
102 Bloor Street, West
Toronto, Ontario, Canada M5S 1M8

J. F. Magosin
Vice President
Systems and Computer Technology
7 North Five Point Road
West Chester, Pennsylvania 19380

Richard S. Mah
Northwestern University
Chemical Engineering Department
Evanston, Illinois 60201

William A. Mahler
Associate Research Psychologist
Educational Testing Service
Princeton, New Jersey 08540

Leslie Maltz
Stevens Institute
Castle Point Station
Hoboken, New Jersey 07030

Michael Marron
Associate Professor of Chemistry
University of Wisconsin-Parkside
Kenosha, Wisconsin 53140

Joseph Matar
Associate Provost
Marquette University
Milwaukee, Wisconsin 53233

Ronald McAdams
Assistant Director
Office for Computing Activities
University of Dayton
Dayton, Ohio 45469

James McDonald
144 Alexander Street
Princeton, New Jersey 08540

John McGeachie
Director of Data Processing
Kiewit Computation Center
Dartmouth College
Hanover, New Hampshire 03755

John P. McGowan
Northwestern University
Evanston, Illinois 60201

George W. McGurn
Director of Research
International Council for Educational
Development
680 Fifth Avenue
New York, New York 10019

David McIntyre
President, Illinois Educational
Consortium for Computer Services
162 Engineering Research Laboratory
University of Illinois
Urbana, Illinois 61801

Ralph McKay
Manager, Financial Information Systems
University of Houston
Cullen Boulevard
Houston, Texas 77004

John P. Menard
Director, Computer Science Center
University of Maryland
College Park, Maryland 20742

Richard L. Merhar
Director, Management Systems - UMIS
University of Pennsylvania
Philadelphia, Pennsylvania 19174

Thomas E. Michels
Assistant Director of the Computer
Network
University of Nebraska
Systems Building
3835 Holdrege, Room 214
Lincoln, Nebraska 68503

E. P. Miles, Jr.
Professor of Mathematics
Florida State University
Tallahassee, Florida

Eugene Miller
New York University
1395 Lexington Avenue
New York, New York 10028

James Miller
President
University of Louisville
Louisville, Kentucky 40208

Russell C. Mills
Associate Vice Chancellor
University of Kansas Medical Center
39th and Rainbow Boulevard
Kansas City, Kansas 66103

Don Mittleman
Director, Computer Center
Oberlin College
Oberlin, Ohio 44074

David C. Montgomery
Director of Institutional Research
and Planning
Oberlin College
Oberlin, Ohio 44074

Roderick Montgomery
Princeton University
Princeton, New Jersey 08540

James L. Morgan
Director, Management Information Systems
State University System of Florida
107 West Gaines Street
Tallahassee, Florida 32304

Alvin L. Morris
Vice President for Administration
University of Kentucky
Lexington, Kentucky 40506

Richard D. Morse
Director, Management Evaluation
and Services
Rutgers, The State University
New Brunswick, New Jersey 08903

Gilbert G. Moser
Executive Director
Educational Informatin Services, Inc.
Hill Center for the Mathematical Sciences
Post Office Box 390
New Brunswick, New Jersey 08903

Charles Mosmann
Computing Facility
University of California
Irvine, California 92664

Wayne A. Muth
Director – Computer Center
West Virginia University
Morgantown, West Virginia 26506

Albrecht J. Neumann
Computer Specialist
National Bureau of Standards
Washington, D.C. 20234

John E. Newton
Instructional Materials Development
Officer
United States Military Academy
West Point, New York 10996

Miguel A. Nogueras
Director, Office of Information Systems
University of Puerto Rico
Post Office Box 4984-G
San Juan, Puerto Rico 00936

William Nugent
Library of Congress
1448 Yellowwood Court
Reston, Virginia 22090

David F. Nyman
Associate Director, Data Systems
Illinois Borad of Higher Education
Springfield, Illinois 62701

William J. O'Brien
Associate Director Inst. Research
Farleigh Dickinson University
1154 River Road
Teaneck, New Jersey 07666

Daniel J. O'Connell
Systems Management Officer
U. S. Naval Academy
Annapolis, Maryland 21402

K. Okashimo
Director, Institute of Computer Science
University of Guelph
Guelph, Ontario, Canada

Thomas Palka
Honeywell Information Systems
1615 Northern Boulevard
Manhasset, New York 11030

Louis T. Parker, Jr
Director, N. C. Educational Computing
Service
The University of North Carolina
Post Office Box 12175
Research Triangle Park,
North Carolina 27709

John R. Pasta
Head, Division of Computer Research
National Science Foundation
Washington, D.C. 20550

Peter Paulson
Director
New York State Library
Albany, New York 12204

Richard T. Paumen
Assistant Vice President – UMIS
University of Pennsylvania
3451 Walnut Street
Philadelphia, Pennsylvania 19174

Girard Pessis
Associate Professor
University of Alberta
Edmonton, Alberta, Canada T6H 1B9

Erving S. Pfau
Assistant Director for Educational
Services
Boston University Computer Center
111 Cummington Street
Boston, Massachusetts 02215

S. J. Plice
Director of Administrative
Systems Planning
University of Michigan
Ann Arbor, Michigan 48104

James Poage
Princeton University
Computer Center
Princeton, New Jersey 08540

Howell R. Porter
Assistant Professor
University of Louisville
Louisville, Kentucky 40208

Icarus Pyros
Head, Office of Computer Science
U. S. Merchant Marine Academy
Kings Point, New York 11024

Joseph Raben
Professor of English
Queens College
Flushing, New York 11367

James Randall
Department of Music
Princeton University
Princeton, New Jersey 08540

Charles B. Ray
Director, Willis H. Booth
Computing Center
California Institute of Technology
Pasadena, California 91109

Joseph B. Reid
Coordonnateur des programmes
Vice-presidence aux communications
Universite du Quebec
2875 Boulevard Laurier
Ste-Foy, Quebec, Canada G1V 2M3

Keith S. Reid-Green
Senior Staff Systems Analyst
Educational Testing Service
Princeton, New Jersey 08540

Roderick R. Ricard, Jr.
Director, Administrative Systems
and Computational Services
University of New Hampshire
Thompson, Durham, New Hampshire 03824

Michael M. Roberts
Deputy Director, Stanford Center
for Information Processing
Stanford University
Stanford, California 94305

Roderick A. Robertson
Assistant Director, Office
Data Processing Management
New Jersey Dept. of Higher Education
10 Bruce Park Drive
Trenton, New Jersey 08618

Stewart H. Robinovitz
ICPR
University of Michigan
Ann Arbor, Michigan 48104

Robert J. Robinson
Computing Center
State University at Albany
1400 Washington Avenue
Albany, New York 12222

Peter Roll
Special Assistant to the Vice President
Academic Administration
University of Minnesota
217 Morrill Hall
Minneapolis, Minnesota 55455

Michael T. Romano
2034 Hart Road
Lexington, Kentucky 40502

Monique Rona
Assistant Director
Computer Center FC-10
University of Washington
Seattle, Washington 98195

Leo Roomets
Director of Computer Services
Massachusetts State College
Computer Network
65 Franklin Street
Boston, Massachusetts 02110

Saul Rosen
Director, Computing Center
Purdue University
West Lafayette, Indiana 47907

J. Michael Ross
Cambridge Project
Massachusetts Institute of Technology
Cambridge, Massachusetts 02138

Judith Rowe
Princeton University
Computer Center
Princeton, New Jersey 08540

John W. Rudan
Director, Office of Computer Services
Cornell University
Ithaca, New York 14850

Dale Rummer
Deputy Director
Computation Center
University of Kansas
Lawrence, Kansas 66044

Jack Sachleben
Project Director
Grumman Data Systems
Bethpage, New York 11714

Gerald St. Amand
Coordinator Computer Programs
College of Business
Michigan State University
East Lansing, Michigan 48824

Ward Sangren
Coordinator of Computer Activities
University of California
2200 University Avenue
Berkeley, California 94720

Anthony P. Santangelo
Director, University Studies
Rutgers, The State University
New Brunswick, New Jersey 08903

August E. Sapega
Trinity College
Hartford, Connecticut 06106

Regina Scarborough
EDUCOM
Princeton, New Jersey 08540

Thomas Scharle
Manager, Academic Applications,
Customer Services, Computing Center
University of Notre Dame
Notre Dame, Indiana 46556

Jeffrey W. Schram
Director, Academic Services
Educational Information Services, Inc.
Hill Center for the Mathematical Sciences
Rutgers, The State University
New Brunswick, New Jersey 08903

Jeffrey A. Schultz
Case Western Reserve University
Department of Operations Research
Cleveland, Ohio 44106

J. D. Seader
University of Utah
Salt Lake City, Utah 84112

Warren D. Seider
Associate Professor, Chemical
Engineering
376 Towne Building
University of Pennsylvania
Philadelphia, Pennsylvania 19104

R. G. Selfridge
Department of Engineering
University of Florida
Gainesville, Florida 32601

Russell Shank
Director of Libraries
Smithsonian Institution
24NHB
Washington, D.C. 20560

Michael S. Sher
Associate Director Center for
Advanced Computation
University of Illinois at Urbana-
Champaign
115 Advanced Computation Building
Urbana, Illinois 61801

Paul Sherburne
Bloomfield College
Bloomfield, New Jersey 07003

Gordon R. Sherman
Director, Computing Center
Tennessee Higher Education Commission
University of Tennessee
Knoxville, Tennessee 37916

Ralph M. Shoffner
Vice President
Richard Abel and Company
Post Office Box 4245
Portland, Oregon 97208

Kenneth Shostack
Office for Information Technology
Harvard University
1730 Cambridge Street
Cambridge, Massachusetts 02138

Paul W. Sire
Director, Office of Management
Information and Computing
University of Vermont
Burlington, Vermont 05401

Patricia C. Skarulis
Acting Assistant Director
Computer Center
Rutgers, The State University
Hill Center
New Brunswick, New Jersey 08904

John E. Skelton
Assistant Professor Mathematics
University of Denver
Denver, Colorado 80210

William D. Slysz
Director of Technical Services
Social Science Data Center
University of Connecticut
Storrs, Connecticut 06268

Jeffry Smith
Systems Head
Educational Testing Service
Princeton, New Jersey 08540

John B. Smith
Professor, Department of English
Pennsylvania State University
University Park, Pennsylvania 16802

Peter J. Smith
DECsystem-10 Educationa Market Manager
Digital Equipment Corporation
146 Main Street
Maynard, Massachusetts 01754

Robin Spock
Baruch College of CUNY
17 Lexington Avenue
New York, New York 10010

Fred Spurgat
Director, Computer Center
Concordia Teachers College
River Forest, Illinois 60305

Einar Stefferud
Einar Stefferud & Associates
3002 Midvale Avenue
Los Angeles, California 90034

Robert Steinhart
Senior Marketing Representativ
IBM Corporation
673 Morris Avenue
Springfield, New Jersey 07081

Paul Stephan
Case Western Reserve University
Dept. of OR. Sears 566
Cleveland, Ohio 44106

Virginia Sterba
New York University
81 Perry Street
New York, New York 10014

Charles D. Sternberg
Computer Systems Analysis
School of Systems and Logistics
Air Force Institute of Technology
Wright-Patterson Air Force Base,
Ohio 45433

Richard Storer
Acting Director Computer Center
Rutgers, The State University
New Brunswick, New Jersey 08903

John C. Storlie
Computer Center Director
University of Wisconsin
La Crosse, Wisconsin 54601

Robert M. Strong
Beth Israel Hospital
Kirstein Building, Room 366
330 Brookline Avenue
Boston, Massachusetts 02215

George Struble
Director, Computing Center
University of Oregon
Eugene, Oregon 97403

Carl F. Stuehrk
Director – Service Plans
American Telephone and Telegraph
195 Broadway
New York, New York 10007

George J. Sullivan
Manager of Computer Services
Rutgers University
Alexander Library
College Avenue
New Brunswick, New Jersey 08901

Lewis J. Sunderland
City University of New York
535 East 80th Street
New York, New York 10021

Len Swanson
Associate Project Director
EDUCOM
Princeton, New Jersey 08540

Lee Talbert
President
Packet Communications, Inc.
176 Second Avenue
Waltham, Massachusetts 02154

Bruce L. Taylor
Director, EISD Division
Educational Testing Service
Princeton, New Jersey 08540

Robert F. Teitel
Technical Manager, Computer Services
The Urban Institute
2100 M Street, N.W.
Washington, D.C. 20037

Karl C. Thomas
Director, Computer Center
Marshall University
Huntington, West Virginia 25701

Thomas R. Thomas
Director of Academic Computing Center
University of Vermont
Burlington, Vermont 05401

Mehran Thomson, Jr.
Director, Exact Sciences
Washtenaw Community College
Post Office Drawer WWC
Ann Arbor, Michigan 48106

Richard S. Touma
Assistant Vice President for
Academic Services
Fairleigh Dickinson University
Rutherford, New Jersey 07070

Cheryl M. Traver
Stanford University
Room 30, Encina Hall
Stanford, California 94305

Giuseppe Traversa
Prof. Matematica Applicata
Universita Degli Studi
L'Aquila, Italy

Larry E. Travis
Director, Academic Computing Center
University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53701

Patricia Treat
EDUCOM
Princeton, New Jersey 08540

Thomas D. Truitt
Board of Directors
Educational Information Services, Inc.
New Jersey Dept. of Higher Education
225 West State Street
Trenton, New Jersey 08625

William Turnbull
President
Educational Testing Service
Princeton, New Jersey 08540

Janet Tuttle
Assistant to the Director
Information Processing Services
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

R. E. Utman
Princeton University Library
P. O. Box 200
Princeton, New Jersey 08540

Dacia Van Antwerp
11111 French Road
Chrysler Institute
Detroit, Michigan 48234

William Van Hassel
Office of Data Analysis Research
Educational Testing Service
Princeton, New Jersey 08540

Roger N. vanNorton
Assistant Deputy Chancellor
New York University
70 Washington Square S.
New York, New York 10003

Henry G. Vaughan
Director, Management Systems
and Analysis
Cornell University
217 Rand Hall
Ithaca, New York 14850

Robert C. Verderber
Office of Science Information Service
National Science Foundation
Washington, D.C. 20550

Bernardo Wacholder
Systems Coordinator – Computer Services
Massachusetts State College
65 Franklin Street
Boston, Massachusetts 02110

Steven V. F. Waite
Faculty Application Resource Adv.
Dartmouth College
Kiewit Computation Center
Hanover, New Hampshire 03755

Harold B. Wakefield
Assistant Chancellor
State University of New York
99 Washington Avenue
Albany, New York 12210

John E. Walker, Director
Computing Center
Lehigh University
Room 119, Packard Lab
Bethlehem, Pennsylvania 18015

Philip Wallack
New York University School of
Education
Higher Education Program
Jersey City State College
Jersey City, New Jersey 07305

R. Brian Walsh
Director, Computing Center
University of Notre Dame
Notre Dame, Indiana 46556

Glenn O. Ware
Assistant Professor
University of Georgia
Office of Computing Activities
Boyd Graduate Building
Athens, Georgia 30602

Charles H. Warlick
Director, Computation Center
University of Texas
Austin, Texas 78712

Arthur Warner
Project Director
Grumman Data Systems
Bethpage, New York 11714

David M. Wax
Acting Director, Northeast Academic
Science Information Center (NASIC)
New England Board of Higher Education
40 Grove Street
Wellesley, Massachusetts 02181

George Weathersby
Associate Director
National Commission on the Financing
of Postsecondary Education
1030 15th Street, N.W., Suite 1060
Washington, D.C. 20005

Gerard P. Weeg
Director, Computer Center
University of Iowa
Iowa City, Iowa 52242

Edward C. Weiss
Program Director, Research Program
Office of Science Information Service
National Science Foundation
Washington, D.C. 20550

Arthur W. Westerberg
University of Florida
Gainesville, Florida 32601

Paul Wesolowski
Director, Management Information and
Institutional Research
Polk Community College
Winter Haven, Florida 33880

Alvin H. White
7280 Church Street
Gilroy, California 95020

V. E. Wichert
Director, Systems
Educational Testing Service
Princeton, New Jersey 08540

David M. Wilson
Assistant Director of System Operations
and Development
Indiana Higher Education Telecommunications
System
1100 West Michigan
Indianapolis, Indiana 46202

David E. Winkel
Director, Division of Computer Services
University of Wyoming
Laramie, Wyoming 82071

William K. Winters
Director of Information Systems
Oklahoma State Regents for
Higher Education
Oklahoma City, Oklahoma 73105

Todd Wipke
Department of Chemistry
Princeton University
Princeton, New Jersey 08540

Joe B. Wyatt
Director, Office for Information Technology
Harvard University
1730 Cambridge Street
Cambridge, Massachusetts 02138

Douwe Yntema
Director, Cambridge Project
Massachusetts Institute of Technology
Cambridge, Massachusetts 02138

Eugene P. Young
Assistant Vice President for Academic Affairs
Rutgers, The State University
New Brunswick, New Jersey 08903

Sally Zeckhauser
Director, Analytic Studies Office
Harvard University
Office for Information Technology
1730 Cambridge Street
Cambridge, Massachusetts 02138

Myron Zimmerman
Student Director of Operations
Juniata College
Huntington, Pennsylvania 16652

Pranas Zunde
Professor
Georgia Institute of Technology
School of Information and Computer Science
Atlanta, Georgia 30332